Schriftenreihe der TMF

M. Nonnemacher | D. Nasseh
J. Stausberg




# Datenqualität in der medizinischen Forschung

## Leitlinie zum adaptiven Management von Datenqualität in Kohortenstudien und Registern

2., aktualisierte und erweiterte Auflage

Medizinisch Wissenschaftliche Verlagsgesellschaft

**Schriftenreihe der TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.**

Band 4, 2. Auflage

Medizinisch Wissenschaftliche Verlagsgesellschaft

**Schriftenreihe der TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.**

Band 4, 2. Auflage

M. Nonnemacher | D. Nasseh | J. Stausberg

# Datenqualität in der medizinischen Forschung

## Leitlinie zum adaptiven Management von Datenqualität in Kohortenstudien und Registern

**2., aktualisierte und erweiterte Auflage**

**unter Mitwirkung von**

U. Bauer | R. Pritzkuleit | C.O. Schmidt | T. Schrader



Medizinisch Wissenschaftliche Verlagsgesellschaft

## Die Autoren

**Dr. rer. medic. Dipl.-Inf. Michael Nonnemacher**
Institut für Medizinische Informatik, Biometrie und Epidemiologie (IMIBE)
Universitätsklinikum Essen
Hufelandstr. 55
45122 Essen

**Prof. Dr. med. Jürgen Stausberg**
Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie (IBE)
Ludwig-Maximilians-Universität München
Marchioninistr. 15
81377 München

**Dipl.-Bioinf. Daniel Nasseh**
Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie (IBE)
Ludwig-Maximilians-Universität München
Marchioninistr. 15
81377 München

MWV Medizinisch Wissenschaftliche Verlagsgesellschaft mbH & Co. KG
Unterbaumstr. 4
10117 Berlin
www.mwv-berlin.de

ISBN 978-3-95466-374-3 (eBook: PDF)

**Bibliografische Information der Deutschen Nationalbibliothek**
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Informationen sind im Internet über http://dnb.d-nb.de abrufbar.



Produkt-/Projektmanagement: Frauke Budig, Berlin
Layout & Satz: eScriptum GmbH & Co. KG – Digital Solutions, Berlin
Herstellung: zweiband.media Agentur für Mediengestaltung und -produktion GmbH

Zuschriften und Kritik an:
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft mbH & Co. KG, Unterbaumstr. 4, 10117 Berlin, lektorat@mwv-berlin.de

## Editorial der TMF

Zentrale Fragen zur Datenqualität im Gesundheitswesen beschäftigen nicht nur Wissenschaftler und Fachverbände seit Jahren, sie rücken zunehmend auch ins Bewusstsein der Gesundheitspolitik und der interessierten Öffentlichkeit. Wie „gut“ sind eigentlich Daten in pharmazeutischen Studien? Wie lässt sich die Datenqualität in Kohortenstudien, Registern und anderen *Data Repositories* messen und gewährleisten, insbesondere wenn diese Daten auch zur Nutzenbewertung medizinischer Leistungen und Produkte herangezogen werden sollen? Welche Datenqualität kann in der Versorgungsforschung und bei der Sekundärdatenanalyse erreicht werden? Mit welchen wissenschaftlichen Methoden und Verfahren lässt sich Datenqualität in der Qualitätsforschung selbst beurteilen und sicherstellen?

Diese Fragen beleuchten schlaglichtartig die wachsenden Herausforderungen an die Beurteilung der Qualität jener Daten, die Zulassungsstudien, Nutzenbewertungen und epidemiologischen Analysen zugrunde liegen und die für die öffentliche Gesundheitssteuerung ebenso unerlässlich sind wie für medizinische *Decision Support*-Systeme. Vor dem Hintergrund der zunehmend digitalen Datenerfassung im Gesundheitswesen ergibt sich die Relevanz dieser Fragen nicht zuletzt aus der Erwartung, dass sich medizinische Daten grundsätzlich auch außerhalb ihres ursprünglichen Kontexts für vielfältige Zwecke nutzen lassen.

Aber was ist eigentlich „Datenqualität“ und wie misst man sie? Ein im vorliegenden Buch hierfür bemühtes Zitat definiert Datenqualität als den Grad, zu dem ein Satz inhärenter Merkmale festgelegte Anforderungen erfüllt. Doch welche Anforderungen sind an medizinische Daten zu stellen – und wie legt man diese fest? Welche Qualitätsindikatoren sind dabei zu berücksichtigen? Und nicht zuletzt stellt sich die Frage, wie sich die Qualität medizinischer Daten schon zum Zeitpunkt ihrer Erhebung positiv beeinflussen lässt. Welcher Aufwand ist damit verbunden und welche Anreizsysteme können langfristig den Aufbau qualitativ hochwertiger Datenbestände für die klinische und Versorgungsforschung vorantreiben?

Als erster gangbarer Schritt in Richtung einer umfassenden Bewertung und nachhaltigen Verbesserung der Datenqualität böte sich auch für Kohortenstudien und wissenschaftliche Register die in klinischen Studien zur Arzneimittelzulassung etablierte *Source Data Verification* an. Ein solch manueller Abgleich zwischen Forschungsdaten und Primärdokumentation (meistens der Krankenakte) erlaubt in der Tat die Bereinigung der meisten Übertragungsfehler auf dem Weg von der Primär- zur Studiendokumentation. Allerdings wären solche Verfahren in großen bevölkerungsbezogenen Studien, wie sie in vielen medizinischen Forschungsverbünden verankert sind, schon allein aus finanziellen und organisatorischen Gründen kaum durchführbar. Alternativ scheint sich hier ein sogenanntes „adaptives Monitoring“, d.h. ein Mix verschiedener Methoden einschließlich Plausibilitätsprüfung und *Source Data Verification*, als

probates Mittel anzubieten, um nicht nur die in Registern und Kohortenstudien schwierige Bewertung der Datenqualität zu bewerkstelligen, sondern um darüber hinaus auch eine kontinuierliche Verbesserung der Arbeit der datenerhebenden Stellen zu erreichen.

Die TMF hat zu diesem für die Epidemiologie, die klinische und die Versorgungsforschung gleichermaßen wichtigen Thema schon vor Jahren einen wertvollen Beitrag leisten können in Form der *Leitlinie zum adaptiven Management von Datenqualität in Kohortenstudien und Registern*. Diese Leitlinie wurde zwischen 2004 und 2006 unter Federführung von Prof. Dr. Jürgen Stausberg (Universitätsklinikum Essen/Kompetenznetz HIV/AIDS, ab 2008: Ludwig-Maximilians-Universität München/Kompetenznetz Hepatitis) in einem TMF-finanzierten Projekt (V020-01) erarbeitet und 2008 in der TMF-Schriftenreihe publiziert.

Wie die Autoren in der Einleitung zum vorliegenden Band ausführlich erläutern, ergab sich schon bald nach ihrer Veröffentlichung ein erster Bedarf für die Überarbeitung der Leitlinie. In zwei von der TMF geförderten Projektabschnitten wurde die alte Leitlinie daher von 2011 bis 2012 zunächst evaluiert (Projekt V020-04) und anschließend einer umfassenden Revision unterzogen (Projekt V020-05). Im Ergebnis kann nun die zweite Version der Leitlinie zur Datenqualität vorgelegt werden. Deren wesentliche Neuerungen bestehen in

- einer deutlichen Erweiterung der Indikatoren zur Messung von Datenqualität,
- einer Aktualisierung der Literaturübersicht und
- spezifischen Anwendungsempfehlungen für Register, Kohorten und andere *Data Repositories*.

Wir freuen uns, diesen wichtigen methodischen Beitrag zeitnah und aktuell in der TMF-Schriftenreihe veröffentlichen zu können. Der vorliegende Band umfasst neben der eigentlichen Leitlinie auch

- eine umfassende und kommentierte Literaturübersicht sowie die Ergebnisse einer Expertenbefragung,
- vielfältige Begriffsdefinitionen,
- Anwendungsempfehlungen für Register, Kohorten und andere *Data Repositories* sowie
- eine Kurzbeschreibung relevanter statistischer Tabellen (Konfidenzintervalle, Fallzahlen), die entweder von der Webseite der TMF heruntergeladen oder bei der TMF-Geschäftsstelle in digitaler Form angefordert werden können.

Die Materialien zur nunmehr abgelösten ersten Version der Leitlinie können bei Bedarf ebenfalls von der TMF bezogen werden (hierzu gehören die Literaturübersicht und das Ergebnis der Expertenbefragung von 2005/06, die auf der ersten Version basierenden Checklisten und Implementierungshilfen sowie die Software-Werkzeuge zur praktischen Nutzung der Leitlinie).

Den Autoren der Leitlinie – Prof. Dr. med. Jürgen Stausberg, Dr. rer. medic. Dipl.-Inf. Michael Nonnemacher und Dipl.-Bioinf. Daniel Nasseh – gilt unser großer Dank für ihre hervorragende Arbeit und für ihr unermüdliches Engagement. Wir sind insbesondere froh über ihren Anspruch, die praktische Anwendung der ursprünglichen Leitlinie zu begleiten und diese durch eine daraus abgeleitete Überarbeitung zu verbessern und besser nutzbar zu machen.

Unser Dank gilt Frau Dr. Ulrike Bauer (Kompetenznetz Angeborene Herzfehler, DHZB), Herrn Prof. Dr. Thomas Schrader (Fachhochschule Brandenburg, Fachbereich Informatik & Medien) und Herrn Dr. Carsten Oliver Schmidt (Institut für Community Medicine-SHIP/KEF, Greifswald) für ihre Beiträge zu den Anwendungsempfehlungen, Herrn Dr. Ron Pritzkuleit (Institut für Krebsepidemiologie e.V., Universität Lübeck) für seine Mitwirkung an der Evaluation der Leitlinie sowie Frau Dorothea Weiland (Institut für Medizinische Informatik, Biometrie und Epidemiologie am Universitätsklinikum Essen), die an der Erstellung der ersten Version maßgeblich beteiligt war.

Die TMF dankt außerdem allen Mitgliedern ihrer Arbeitsgruppe „IT-Infrastruktur und Qualitätsmanagement“, die unter der Leitung von Ronald Speer (IMISE und KKS Universität Leipzig) bzw. Prof. Dr. Ulrich Sax (Universitätsmedizin Göttingen) kontinuierlich die Evaluation der Leitlinie und ihre Überarbeitung begleitet und sich um ein kritisches Feedback der Anwender in den medizinischen Forschungsverbünden bemüht haben. Aus diesem Kreis ist insbesondere Gisela Antony (Kompetenznetz Parkinson, Universität Marburg) hervorzuheben, die viele hilfreiche Hinweise zur praktischen Anwendung der Leitlinie gegeben hat.

Mit der vorliegenden Leitlinie hofft die TMF, gleichermaßen Hilfestellung und Denkanstoß zum wichtigen Thema „Qualität medizinischer Forschungsdaten“ zu liefern. Da sich die daran geknüpften methodischen Fragen nie abschließend beantworten lassen werden, wird sich die TMF auch weiterhin auf dem Gebiet engagieren. Jedes Feedback zur Leitlinie ist den Mitgliedern unserer AG „IT-Infrastruktur und Qualitätsmanagement“ daher herzlich willkommen!

Für die TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. (TMF) im Auftrag des Vorstands

*Sebastian Claudius Semler*
(Geschäftsführer)

*Prof. Dr. Michael Krawczak*
(Vorstandsvorsitzender)

# Inhalt

# I Einleitung

Der Nutzen krankheitsbezogener wissenschaftlicher Kohortenstudien und Register, wie sie von Kompetenznetzen in der Medizin, Deutschen Zentren der Gesundheitsforschung und zahlreichen anderen medizinischen Forschungsnetzwerken aufgebaut wurden und werden, hängt ganz wesentlich von der Qualität der dort erfassten Daten ab [Gliklich/Dreyer 2010, Müller et al. 2010]. Datenqualität wird vom Deutschen Institut für Normung in der DIN EN ISO 14050:2010-08 definiert als „Eigenschaften von Daten in Bezug auf ihre Eignung, festgelegte Anforderungen zu erfüllen" [DIN EN ISO 14050 2010]. Zur effizienten Sicherstellung einer hohen Datenqualität sind wissenschaftlich abgesicherte Methoden und Verfahren vonnöten. Die vorliegende Leitlinie enthält daher Empfehlungen zum Management der Datenqualität von Registern, Kohortenstudien und Data Repositories. Sie wendet sich an Verantwortliche für das Datenmanagement und die Qualitätssicherung in Studien- oder Registerzentralen. Hierbei sollte es sich um Personen mit wissenschaftlichem Studium und relevanten Vorerfahrungen handeln. Bei der Anwendung der hier beschriebenen statistischen Verfahren zur Fallzahlplanung (s. Kap. II.5) kann die Hinzuziehung eines Statistikers oder Biometrikers sinnvoll sein.

Die Erstellung von Version 1.0 dieser Leitlinie vom 31.3.2006 wurde durch die TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. im Rahmen des Projektes „Entwicklung, Umsetzung und Evaluation von Leitlinien zum adaptiven Management von Datenqualität in Kohortenstudien und Registern. Teil 1: Entwicklung und Umsetzung der Leitlinien" gefördert. Im Jahre 2011 führten dann zwei Entwicklungen zu einer erneuten Auseinandersetzung mit dieser Version. Zum einen wurde in zwei Veröffentlichungen über eine erfolgreiche Anwendung der Leitlinie berichtet [Jacke et al. 2010, Prokein et al. 2010]. Zum anderen wurden – auch von TMF-Mitgliedsverbünden – komplementäre Ansätze zur Operationalisierung von Datenqualität vorgestellt, so dass die Aktualität der Leitlinie in Version 1.0 in Frage stand. Ein Workshop der TMF im April 2011 in Berlin bestätigte dann den

Bedarf einer Überarbeitung[1]. Dort wurde in den Themenbereichen Methoden, Datenqualität in der Forschung und Datenqualität in der Versorgung eine Bestandsaufnahme vorgenommen. In dieser Situation förderte die TMF in den Jahren 2011 und 2012 ein Projekt, bei dem Überlegungen aus epidemiologischen Krebsregistern, einer Kohortenstudie sowie einem Data Repository bezüglich der Messung von Datenqualität mit der Version 1.0 der Leitlinie abgeglichen wurden[2]. Als Ergebnis wurde deren konzeptioneller Ansatz bestätigt, das Set von Indikatoren zur Messung der Datenqualität jedoch deutlich erweitert [Stausberg et al. 2012].

Die Einbindung neuer Indikatoren stellt daher die wesentliche Änderung in Version 2.0 der Leitlinie dar. Weitere Teile der Leitlinie wurden entsprechend angepasst; die Literatursichtung wurde bis zum Jahr 2013 aktualisiert. Eine Erweiterung der Leitlinie findet sich mit Anwendungsempfehlungen zu Registern, Kohorten und Data Repositories. Die Erweiterung ergibt sich aus der Verbreiterung des Indikatorensets, bei dem in noch höherem Maß als in Version 1.0 ein Zuschnitt auf die Gegebenheiten des jeweiligen Vorhabens sinnvoll bzw. notwendig ist. Die Anwendungsempfehlungen sind dabei als beispielhafte Muster für jeweils einen Typus von empirischen Forschungsvorhaben zu verstehen. In Gliederung und Darstellung orientiert sich Version 2.0 der Leitlinie an der ersten Fassung. Damit wird Nutzern, die bereits mit Version 1.0 vertraut sind, ein reibungsloser Umstieg ermöglicht. Zur Abwärtskompatibilität wurden in Einzelfällen Gliederungspunkte ohne Inhalt belassen. Auch die Revision der Leitlinie wurde durch die TMF ideell und finanziell gefördert.

Die Leitlinie bindet sich in ein vielfältiges Instrumentarium sinnvoller Maßnahmen im Datenmanagement von Kohortenstudien und Registern ein. Tabelle 1 zeigt ein Rahmenmodell, welches hierzu von Arts et al. [2002] erarbeitet wurde. In diesem Vorhaben wurden gezielt zwei Maßnahmen herausgegriffen, die sich im Umfeld der TMF als besonders relevant herauskristallisiert haben, ohne dass hierfür jedoch bereits klare Verfahrensanweisungen gegeben werden können. Dies sind der Originaldatenabgleich (Source Data Verification, SDV) und das Feedback.

Es wird davon ausgegangen, dass eine komplette Überprüfung der gemeldeten Daten mit der beim Datenlieferanten vorliegenden Primärdokumentation, also eine komplette SDV, zwar wünschenswert, aber nicht umsetzbar ist. Dies liegt zum einen an der fehlenden Finanzierbarkeit, zum anderen aber auch an der Unpraktikabilität dieses Vorgehens. Schwerpunkt der Leitlinie ist daher die Definition eines adaptiven Konzeptes zum Monitoring, welches eine dynamische Anpassung der Aktivitäten zur Qualitätsprüfung an die Datenquali-

1 Programm und Unterlagen der Vortragenden finden sich unter http://www.tmf-ev.de/Themen/Projekte/V020_01_Datenqualitaet.aspx.

2 Der Schlussbericht zu diesem TMF-Projekt (V020-04) ist unter www.tmf-ev.de/Leitlinie-Datenqualitaet zum Download verlinkt.

Tab. 1 Generisches Rahmenmodell zur Datenqualität in Registern aus Arts et al. [2002a]

| Central coordinating centre | Local sites |
|---|---|
| **Prevention during set up and organisation of registry** | |
| **At the onset of the registry**<br>■ compose minimum set of necessary data items<br>■ define data & data characteristics in data dictionary<br>■ draft a data collection protocol<br>■ define pitfalls in data collection<br>■ compose data checks<br>■ create user friendly case record forms<br>■ create quality assurance plan<br>**In case of new participating sites**<br>■ perform site visit<br>■ train new participants<br>**Continuously**<br>■ motivate participants<br>■ communicate with local sites<br>**In case of changes (e.g., in data set)**<br>■ adjust forms, software, data dictionary, protocol, training material, etc.<br>■ communicate with local sites | **At the onset of participating in the registry**<br>■ assign a contact person<br>■ check developed software for data entry and for extraction<br>■ check reliability and completeness of extraction sources<br>■ standardise correction of data items<br>**Continuously**<br>■ train (new) data collectors<br>■ motivate data collectors<br>■ make data definitions available<br>■ place date & initials on completed forms<br>■ keep completed case record forms<br>■ data collection close to the source and as soon as possible<br>■ use the registry data for local purposes<br>**In case of changes (e.g., in data set)**<br>■ adjust data dictionary, forms, software, etc.<br>■ communicate with data collectors |
| **Detection during data collection** | |
| **During import of data into the central database**<br>■ perform automatic data checks<br>**Periodically and in case of new participants**<br>■ perform site visits for data quality audit (registry data < > source data) and review local data collection procedures<br>**Periodically**<br>■ check inter- and intraobserver variability<br>■ perform analyses on the data | **Continuously**<br>■ visually inspect completed forms<br>■ perform automatic data checks<br>■ check completeness of registration |
| **Actions for quality improvement** | |
| **After data import and data checks**<br>■ provide local sites with data quality reports<br>■ control local correction of data errors<br>**After data audit or variability test**<br>■ give feedback of results and recommendations<br>■ resolve causes of data errors | **After receiving quality reports**<br>■ check detected errors<br>■ correct inaccurate data & fill in incomplete data<br>■ resolve causes of data errors<br>**After receiving feedback**<br>■ implement recommended changes<br>■ communicate with personnel |

tät ermöglicht. Diese Anpassung umfasst die Frequenz, den Umfang und die Tiefe der SDV. Die Prüfung der Datenqualität erfolgt dabei an zwei Stellen:

1. Zum einen werden Indikatoren der Datenqualität definiert, die aus den Daten eines Registers oder einer Kohortenstudie direkt erhoben werden können.
2. Zum anderen ist der Grad von Vollzähligkeit, Vollständigkeit und Übereinstimmung zwischen gemeldeten Daten und der vor Ort existierenden Dokumentation ein Maß zur Anpassung des Monitorings.

Ein adaptives Monitoring verfolgt nicht das Ziel, eine vollständige Übereinstimmung der erhobenen Daten mit der Primärdokumentation sicherzustellen. Im Sinne einer Qualitätslenkung wird vielmehr eine Beeinflussung der meldenden Leistungserbringer angestrebt, um über das Konzept der kontinuierlichen Qualitätsverbesserung bereits bei der Erfassung eine hohe Datenqualität zu erzielen. Abbildung 1 verdeutlicht diesen Sachverhalt.

Derzeit gibt es zu einem adaptiven Monitoring weder Vorgaben für Register (z.B. Krebsregister) noch wissenschaftlich abgesicherte Verfahren. Alle medizinischen Forschungsverbünde mit zentraler Datenerhebung, ob als Kohorte, epidemiologische Studie oder Register, stehen daher vor der gleichen Aufgabe. Im Vorlauf zur Erstellung von Version 1.0 der Leitlinie wurde im Februar 2005 in Essen ein Workshop „Medizinische Daten in Forschung und Versorgung“ durchgeführt[3]. Dort wurde einerseits die Notwendigkeit eines adaptiven Monitorings aus ganz unterschiedlichen Sichten (Therapieoptimierungsstudien, wissenschaftliche Register, externe Qualitätssicherung) eindrücklich betont. Anderseits wurden die fehlende konzeptionelle Basis und der Bedarf nach einer Ausarbeitung von Leitlinien und deren konkrete Unterstützung durch Software-Werkzeuge deutlich.

Auch wenn die SDV weiterhin eine wichtige Rolle im Management von Datenqualität einnimmt, gewinnt die Datenerfassung ohne Verfügbarkeit von Originaldaten zunehmend an Bedeutung. Hierbei handelt es sich z.B. um Kohorten, bei denen die Erfassung der Daten ausschließlich zum Zwecke des Vorhabens unmittelbar durch studieneigenes Personal erfolgt. Daten werden auch bei ihrer Erhebung in Interviews oder Untersuchungen unmittelbar in zentralen Systemen erfasst, ohne einen Zwischenschritt über eine lokale, evtl. papierbasierte Dokumentation. Patienten und Angehörige geben selbst Daten zum funktionellen Outcome in zentralen Systemen zum Electronic Data Capture (EDC) ein oder führen dort Aufzeichnungen zu Schmerzen und anderen Symptomen. Daten werden aus medizintechnischen Systemen wie Sensoren automatisch übernommen, so dass das Studiensystem den einzigen permanenten Speicherort darstellt. In diesen Fällen bietet sich alternativ das Feed-

3 Programm, Unterlagen der Vortragenden und eine Zusammenfassung finden sich unter http://www.ekmed.de/routinedaten/.

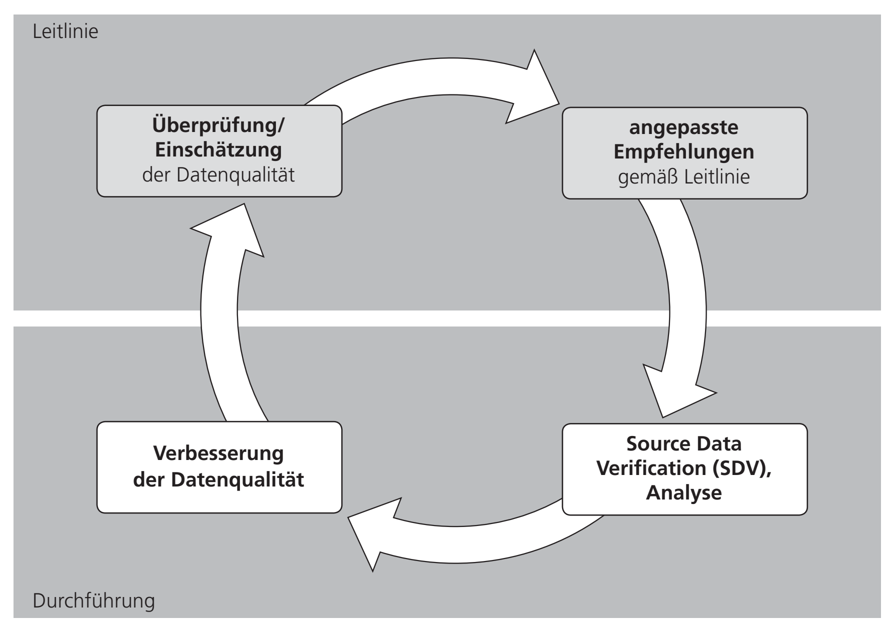



Abb. 1 Kontinuierliche Verbesserung der Datenqualität

back als Maßnahme im Qualitätsmanagement an. Die Messung der Qualität der Daten ist aber auch erforderlich, um ihre Eignung für den jeweiligen Anwendungszweck zu prüfen. Wird hier ein bestimmtes Niveau unterschritten, können Daten zur Beantwortung bestimmter Fragestellungen nicht mehr geeignet sein [Malin/Keating 2005]. Auch für die elektronische Akte in der Versorgung wurden differenzierte Anforderungen an die Datenqualität abhängig vom Anwendungszweck der Dokumentation beschrieben [Baxter et al. 2007].

Der erste Schritt bei der Entwicklung der Leitlinie in Version 1.0 war die Identifikation und Analyse der bestverfügbaren Evidenz über eine Sichtung der Literatur und eine Befragung von Experten[4]. Daran schlossen sich die Konsensusfindung zur inhaltlichen Zusammenführung und Ergänzung der Ergebnisse und die Ableitung der Leitlinie an. Ein Zwischenstand wurde der interessierten Fachöffentlichkeit im Februar 2006 in einem Workshop „Qualität medizinischer Daten in Forschung und Versorgung" präsentiert[5]. Auf diesem Workshop wurden unter anderem in vier Beiträgen Evidenz, Empfehlungen und Anwendung der Leitlinie erläutert und diskutiert. Die Vorgehensweise orientiert sich an Empfehlungen zur Entwicklung von Leitlinien, die national erarbeitet [AWMF/ÄZQ 2001] und wissenschaftlich begründet wurden

4 Die Literatursichtung bis 2005/2006 steht unter www.tmf-ev.de/Leitlinie-Datenqualitaet zum Download zur Verfügung.
5 Programm, Unterlagen der Vortragenden und eine Zusammenfassung finden sich unter http://www.ekmed.de/routinedaten/.

[Bundesärztekammer/Kassenärztliche Bundesvereinigung 2005]. Das Vorgehen bei der Literatursichtung entspricht den Schritten zur Quellensuche und Quellenbewertung, wie es die BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH für die externe vergleichende Qualitätssicherung nach § 137 SGB V beschreibt [Mohr et al. 2005]. Die Fortschreibung der Leitlinie zur Version 2.0 erfolgte hingegen vor dem Hintergrund konkreter Beispiele für das Management von Datenqualität. Die gleichzeitig vorgenommene Aktualisierung der Literaturbasis ist daher komplementär zu verstehen und war nicht Ausgangspunkt zur Einführung der neuen Indikatoren.

Unabhängig von den hier gewonnenen Ergebnissen wurde festgelegt, durchgehend das Glossar der Arbeitsgruppe Qualitätsmanagement in der Medizin der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) als Referenz für die Definition qualitätsrelevanter Begriffe zu Grunde zu legen [GMDS-AG Qualitätsmanagement in der Medizin 2003]. Dieses 1996 erschienene und 2003 sowie 2007 [Sens et al. 2007] grundlegend überarbeitete und aktualisierte Glossar zu Begriffen und Konzepten des Qualitätsmanagements findet sowohl in der Fachöffentlichkeit als auch in der Praxis durchgehend Anerkennung und kann derzeit in Deutschland als Defacto-Standard gelten.

***Hinweis***

***Bei geschlechtsspezifischen Bezeichnungen wird in der Leitlinie nur eine Form genannt. Damit sind beide Geschlechter gemeint, solange nicht ausdrücklich anders vermerkt.***

# II Vorgehensweise

## II.1 Übersicht

Abbildung 2 gibt einen Überblick zur Verwendung der Leitlinie im Datenmanagement in Form eines Aktivitätsdiagramms der Unified Modeling Language (UML).

Wie bei Arts et al. 2002a wird hierbei zwischen dem zentralen Datenmanagement und den einzelnen Erhebungszentren unterschieden. Bereitstellung der erforderlichen Informationen sowie Anwendung der Leitlinie fallen in die Ver-

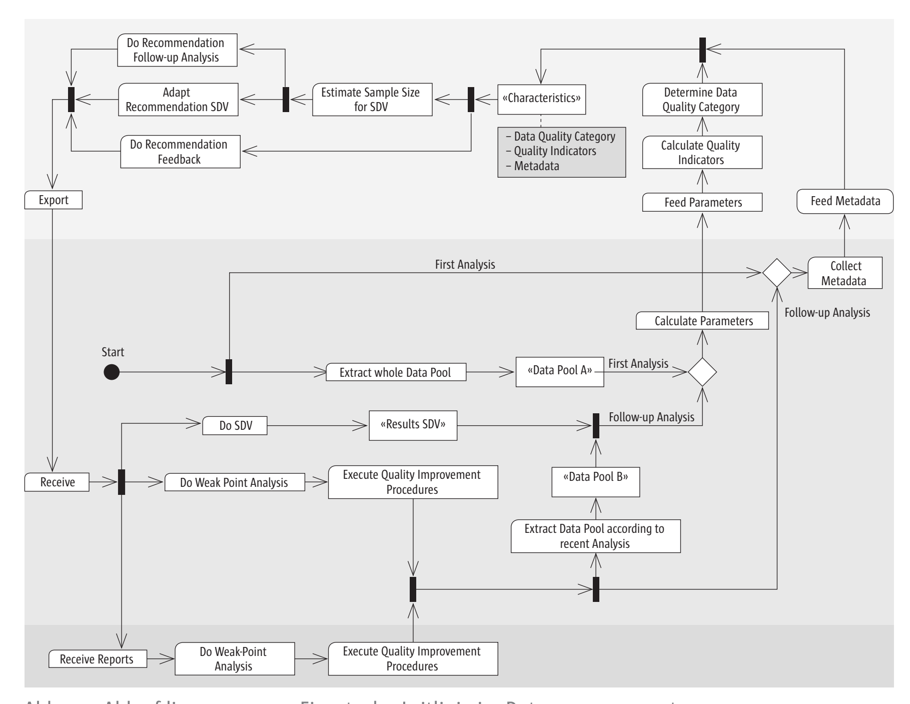



Abb. 2 Ablaufdiagramm zum Einsatz der Leitlinie im Datenmanagement

antwortung des zentralen Datenmanagements. Die einzelnen Erhebungszentren sind mittelbar bei der SDV sowie unmittelbar als Zielgruppe des Berichtswesens betroffen. Der Zyklus zur kontinuierlichen Qualitätsverbesserung ist in Abbildung 2 entgegen dem Uhrzeigersinn dargestellt und gliedert sich in die Teile Vorbereitung, Anwendung der Leitlinie sowie Umsetzung der Empfehlungen. Die involvierten Bereiche sind in horizontalen Zeilen dargestellt, das zentrale Datenmanagement findet sich mittig, die Erhebungszentren am unteren, die zur Anwendung empfohlene aber nicht zwingend erforderliche Software am oberen Rand. Die Besonderheit beim ersten Zyklus liegt in der Nutzung des gesamten Datenpools sowie dem Fehlen der Parameter aus der SDV. Ab dem zweiten Zyklus werden nur noch Meldungen und Beobachtungen berücksichtigt, die seit dem vorhergehenden Durchlauf eingetroffen sind.

Gemäß dem Projektansatz findet eine Adaptation (= Anpassung) allgemeiner guter Praxis im Datenmanagement an Charakteristika konkreter Vorhaben gleich mehrfach statt. Im Mittelpunkt steht eine Anpassung an die Datenqualität, die über verschiedene Qualitätsindikatoren abgeschätzt wird. Daneben fließen Ressourcen (z.B. für SDV), Design (z.B. Zahl an Merkmalen) oder auch Meldezahlen pro Zentrum in die aus der Leitlinie gewonnenen Empfehlungen ein.

## II.1.1 Vorbereitung

Im Rahmen der Vorbereitung fallen regelmäßig drei Aufgaben an (s. Abb. 3):

1. Die Extraktion des Datenbestandes seit der letzten Analyse (bzw. des gesamten Datenbestandes im ersten Zyklus). Dieser Datenbestand wird im Folgenden als Datenpool bezeichnet.
2. Die Zusammenstellung von Metadaten des Vorhabens.
3. Die Ermittlung derjenigen Parameter, die als Zähler oder Nenner für die Berechnungen der Qualitätsindikatoren herangezogen werden.

Es wird empfohlen, diese Aufgaben sequentiell zu bearbeiten.

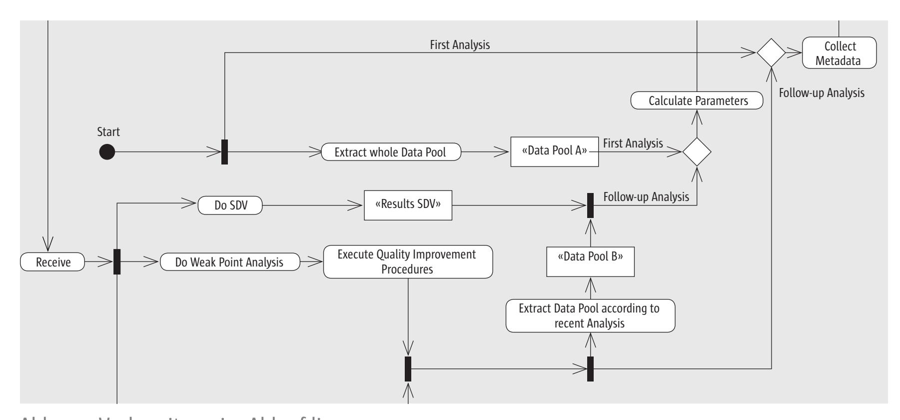



Abb. 3 Vorbereitung im Ablaufdiagramm

Die Extraktion oder Selektion des Datenbestandes fällt in die Verantwortung des zentralen Datenmanagements. Es wird davon ausgegangen, dass im entstehenden Datenbestand Beobachtungseinheiten definiert werden können, deren Erfassung bzw. Übernahme komplett in den Auswertungszeitraum fällt. Als Beobachtungseinheiten können je nach Vorhaben Personen, Patienten, Behandlungsfälle, Meldungen, Erhebungsbögen u.a. bezeichnet sein. Der Vorgang der Extraktion wird in dieser Leitlinie nicht abgehandelt.

Metadaten ermöglichen eine Anpassung der Empfehlungen an konkrete Vorhaben, ohne dass die Metadaten Aussagen zur Datenqualität beinhalten. Metadaten sind bei jedem Zyklus zu erheben bzw. zu aktualisieren. Eine Beschreibung der Metadaten findet sich in Kapitel II.4. Einige Merkmale der Metadaten sind aus dem Datenpool zu ermitteln, andere finden sich im Studienplan oder anderen Dokumenten des Vorhabens.

Für jeden Qualitätsindikator sind zwei Parameter zu ermitteln, die als Zähler bzw. Nenner bei der Ermittlung der Rate herangezogen werden. Für den überwiegenden Teil der Parameter muss bei ihrer Ermittlung auf den Datenpool zugegriffen werden. Die Parameter sind mit Berechnungshinweisen bei den Qualitätsindikatoren in Kapitel II.2 dargestellt. Die Parameter sind für jedes Zentrum getrennt zu erheben.

### II.1.2 Anwendung

Die Anwendung der Leitlinie sollte bevorzugt rechnerunterstützt erfolgen. Die hier dargestellten Schritte laufen dann automatisch ab. Neben den im Rahmen der Vorbereitung ermittelten Parametern (Metadaten, Zähler und Nenner der Qualitätsindikatoren) sind hierzu keine weiteren Eingaben erforderlich. Um die automatische Anwendung zu ermöglichen, müssen verschiedene Festlegungen getroffen werden, z.B. von Schwellenwerten und Gewichten. Auf Grund fehlender Angaben in der Literatur wurde in Version 2.0 der Leitlinie auf die Empfehlung von Schwellenwerten verzichtet; die aufgeführten Gewichte sind beispielhaft zu verstehen. Bei der Anwendung der Leitlinie sind daher Schwellenwerte festzulegen und die beispielhaft genannten Gewichte zu überprüfen. Die Verantwortung hierzu trägt das zentrale Datenmanagement.

Die Anwendung der Leitlinie lässt sich in folgende Schritte zerlegen (s. Abb. 4):

1. Kalkulation der Qualitätsindikatoren
   - Berechnung der Raten
   - Zuordnung der Raten zu den Kategorien auffällig/unauffällig
2. Ermittlung des Qualitätsscores
3. Durchführung der Fallzahlplanung
   - Ermittlung der Zahl zu untersuchender Beobachtungseinheiten (Umfang)

- Ermittlung der Zahl der Merkmale, die pro Beobachtungseinheit zu untersuchen sind (Tiefe)
- Ermittlung der Zahl von Besuchen der Erhebungszentren (Frequenz)

4. Zusammenstellung der Berichte für das Feedback
5. Festlegung des Zeitpunktes für die nächste Analyse

Alle Schritte sind für jedes Zentrum getrennt auszuführen.

Die Raten der Qualitätsindikatoren ergeben sich durch Division von Zähler und Nenner nach den Definitionen in Kapitel II.2.

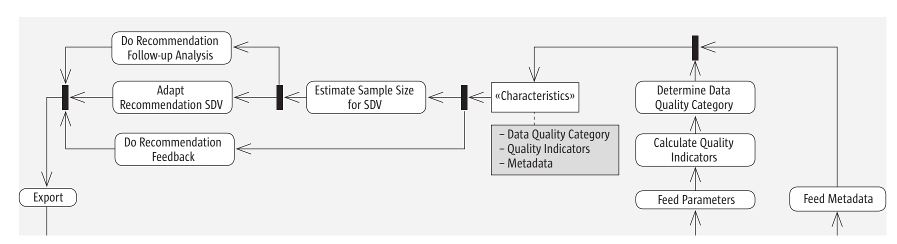



Abb. 4 Anwendung im Ablaufdiagramm

Für die Trennung von unauffälligen (guten) Raten von auffälligen (schlechten) Raten werden keine allgemeinen Empfehlungen für Schwellenwerte gegeben. Diese sind daher vorhabensspezifisch festzulegen. Für einzelne Indikatoren finden sich Angaben zu möglichen Schwellenwerten bei den Anwendungsempfehlungen. Bei Über- oder Unterschreiten der Schwellenwerte ergibt sich eine Bewertung der ermittelten Rate als auffällig oder unauffällig. Für kleinere Fallzahlen wird eine alternative Zuordnung angeboten, bei der der Einschluss des Schwellenwertes in ein 95%-Konfidenzintervall um die gemessene Rate geprüft wird. Bei Einschluss des Schwellenwertes wird – „im Zweifel für das Erhebungszentrum" – der Indikator als unauffällig gewertet.

Die Qualitätsindikatoren werden über eine gewichtete Summe zu einem Score mit Werten zwischen 0 (alle gemessenen Raten sind auffällig) und 100 (alle gemessenen Raten sind unauffällig) verdichtet. Für die Gewichte wurden normativ-analytisch Empfehlungen erarbeitet, die ebenfalls vorhabensspezifisch verändert werden können. Hierzu trifft die Leitlinie keine Regelungen. Eine Auflistung der Gewichte findet sich in Tabelle 3 in Kapitel II.3. Der Score wird weiterhin in fünf Kategorien von Datenqualität unterteilt, welche u.a. bei der Fallzahlplanung den absoluten Abstand der unteren und oberen Grenze des 95%-Konfidenzintervalls zum Schätzer festlegen. Die Ermittlung des Qualitätsscores ist in Kapitel II.3 beschrieben.

Eine Fallzahlplanung für eine SDV ist nur unter zwei Voraussetzungen sinnvoll: Originaldaten stehen zur Verfügung und eine SDV ist möglich. Bei kompletter Anonymisierung von Daten in einem Vorhaben trifft letzteres z.B. nicht zu. In die Fallzahlplanung fließen folgende Parameter ein:

- Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Beobachtungseinheiten (Qualitätsindikator), zentrumsspezifisch. Im ersten Zyklus wird hier der vorab festgelegte Schwellenwert eingesetzt. In den folgenden Zyklen die ermittelte Rate.
- Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Datenelemente (Qualitätsindikator), zentrumsspezifisch. Im ersten Zyklus wird hier der vorab festgelegte Schwellenwert eingesetzt. In den folgenden Zyklen die ermittelte Rate.
- Absoluter Abstand der unteren und oberen Grenze des 95%-Konfidenzintervalls zum Schätzer, einheitlich für alle Zentren. Dieser Abstand wird je nach Kategorie der Datenqualität gewählt. Er ist größer bei besserer Datenqualität und kleiner bei schlechterer Datenqualität. Die empfohlenen Werte finden sich in Kapitel II.5.
- Anzahl der Beobachtungseinheiten im Analysezeitraum, zentrumsspezifisch. Die Werte finden sich in den Metadaten.

Das Vorgehen bei Fallzahlplanung ist in Kapitel II.5 detailliert vorgestellt. Ein Überblick zu den Differenzierungen in Umfang, Tiefe und Frequenz findet sich am Ende dieses Abschnittes.

Entsprechend den Vorgaben in den Metadaten wird zwischen drei Berichtsvarianten ausgewählt. Die Prinzipien des Feedbacks beschreibt Kapitel II.6.

Der Zeitpunkt für die nächste Analyse ist mit einem Default von sechs Monaten belegt. Kapitel II.7 enthält Hinweise auf eine Anpassung an die Gegebenheiten eines Vorhabens.

### II.1.2.1 Ermittlung der Zahl zu untersuchender Beobachtungseinheiten (Umfang)

Zuerst wird für jedes Zentrum die Zahl der Beobachtungseinheiten ermittelt, für die in der folgenden Phase ein Abgleich mit den Originaldaten durchgeführt wird. Die gewählte Formel stellt sicher, dass auch bei kleinen Zentren die empfohlenen Fallzahlen nicht höher liegen als die im vorhergehenden Zyklus aufgetretene Gesamtzahl. Dies wird als Umfang der SDV bezeichnet.

### II.1.2.2 Ermittlung der Zahl der Merkmale, die pro Beobachtungseinheit zu untersuchen sind (Tiefe)

Anschließend wird für jedes Zentrum ermittelt, bei wie vielen Merkmalen einer Beobachtungseinheit die Überprüfung durchgeführt wird. Das Vorgehen ist formal identisch zu Kapitel II.1.2.1 mit folgenden Besonderheiten: Der gemessene Wert entspricht der relativen Häufigkeit von fehlenden Übereinstimmungen bei allen geprüften Beobachtungseinheiten. Die Fallzahlplanung bezieht sich auf die fehlende Übereinstimmung bei einer Beobachtungseinheit. Auch hier ist durch die gewählte Formel sichergestellt, dass die empfohlene Zahl von Merkmalen nicht höher liegt als die Gesamtzahl von Merkmalen

pro Beobachtungseinheit. Es wird davon ausgegangen, dass die Gesamtzahl von Merkmalen eines Vorhabens über alle Zentren identisch ist. Dieser Teil der Fallzahlplanung legt die Tiefe der SDV fest.

### II.1.2.3 Ermittlung der Zahl von Besuchen der Erhebungszentren (Frequenz)

Die in Kapitel II.1.2.1 ermittelte Fallzahl wird durch die in den Metadaten beschriebene maximale Bearbeitungszahl pro Besuch eines Erhebungszentrums dividiert. Hierdurch ergibt sich die Zahl der Besuche für die SDV. Als Dauer eines Zyklus werden sechs Monate empfohlen. Sollte die ermittelte Fallzahl die Bearbeitungskapazität überschreiten, wird dies im Bericht vermerkt und die in den Metadaten angegebene Bearbeitungskapazität auf den Zyklus verteilt. Dieser Teil der Planung legt die Frequenz der Besuche für die SDV fest.

## II.1.3 Ausführung der Empfehlungen

Die Ausführung der Empfehlungen betrifft das zentrale Datenmanagement sowie die einzelnen Erhebungszentren (s. Abb. 5). Weitere Regelungen hierzu trifft die Leitlinie nicht.

Empfänger der Empfehlungen ist das zentrale Datenmanagement (1). Es leitet die Berichte zur Datenqualität an die einzelnen Erhebungszentren weiter (2). Das zentrale Datenmanagement führt die SDV nach den Empfehlungen aus (3). Zentrales Datenmanagement und lokale Erhebungszentren sollten auf Grund der Berichte eine Schwachstellenanalyse durchführen (4) und Maßnahmen zur Verbesserung der Datenqualität ergreifen (5). Damit ist ein Zyklus beendet. Der nächste Zyklus beginnt nach der festgelegten Zeit (Default sechs Monate) mit der Vorbereitung seitens des zentralen Datenmanagements.

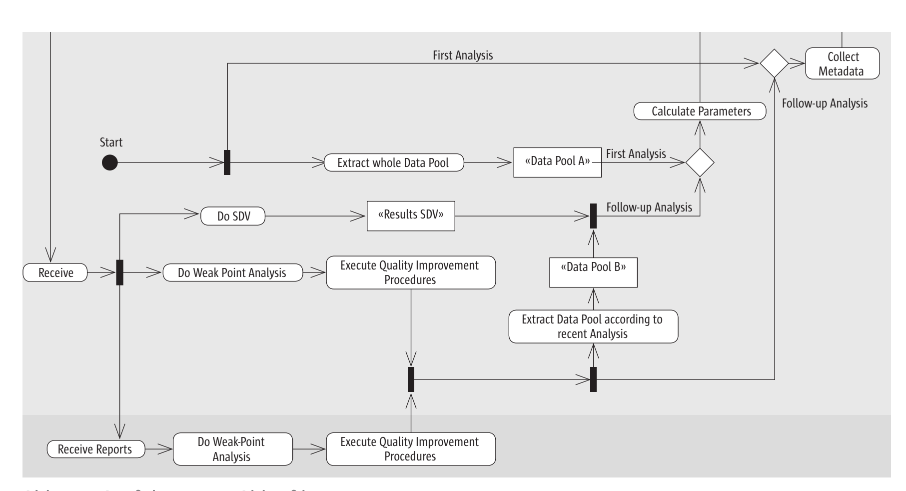



Abb. 5 Ausführung im Ablaufdiagramm



## II.2 Qualitätsindikatoren

Aus Literatursichtung, Expertenbefragungen und Anwendungsverfahren lassen sich eine Vielzahl von Kenngrößen als Indikatoren für die Messung von Datenqualität identifizieren. Diese Qualitätsindikatoren werden in Kategorien zusammengefasst, die den von Donabedian eingeführten Ebenen der Strukturen, Prozesse und Ergebnisse entsprechen [Donabedian 1988]. Diese Ebenen haben sich in der Diskussion um die Qualität der medizinischen Versorgung als wertvoll erwiesen. Das in dieser Leitlinie betrachtete Produkt sind dabei die Daten, die in der Registerzentrale oder der Studienzentrale gespeichert sind. Als Hilfestellung für den Leser werden daher für die von Donabedian geprägten Bezeichnungen Synonyme eingeführt, die aus Sicht des Datenmanagements und Monitorings eher erschließbar sind; diese sind: für Struktur die Integrität, für Prozess die Organisation sowie für Ergebnisse die Richtigkeit.

Zur Beurteilung der Datenqualität (und zur Bildung eines evaluativen Scores, s. Kap. II.3) muss pro Indikator ein individueller Schwellenwert für die Unterscheidung von auffällig/unauffällig festgelegt werden [JCAHO 1990].

In Version 2.0 umfasst die Leitlinie insgesamt 51 Indikatoren, 30 Indikatoren der Ebene Integrität, 15 der Ebene Organisation und 6 der Ebene Richtigkeit. Hiervon findet sich neu ein Indikator zur Qualität von Metadaten, Indikator TMF-1050 „Umfang der Metadaten bei Untersuchungen“. Während dieser im Vorprojekt noch einer neuen Ebene „Metadaten“ zugeordnet war, ist der Indikator nun in die Ebene Integrität eingeordnet. Im Folgenden werden die Indikatoren im Detail vorgestellt. Die Darstellung der Indikatoren ist angelehnt an Empfehlungen der Joint Commission on Accreditation of Healthcare Organizations (JCAHO) für die Beschreibung evidenz-basierter klinischer Qualitätsindikatoren im Gesundheitswesen [JCAHO 1990]. Tabelle 2 zeigt beispielhaft die Bedeutung der einzelnen Elemente der Definition eines Indikators.

Tab. 2 Aufbau einer Indikatordefinition

| |
|---|
| **Beschreibung** |
| z.B. *Anzahl fehlender Werte* |
| **Begriffsdefinitionen** |
| werden auch im Glossar aufgeführt |
| **Identifikationsnummer (ID)** |
| Eindeutiges Kennzeichen jedes Indikators. Die Nummerierung hat keine Bedeutung. |
| **Typ** |
| Strukturqualität/Prozessqualität/Ergebnisqualität |
| **Quellen** |
| Literaturquellen, Expertenmeinung |
| **Bezug** |
| Berechnung für einzelne Datenelemente oder für größere Informationseinheiten (z.B. Meldung, Register) |
| **Alternative Definition** |
| Hinweis auf alternative Definitionen |
| **Anmerkungen** |
| |
| **Zähler** |
| z.B. *Anzahl fehlender Werte des Datenelements* |
| **Nenner** |
| z.B. *Anzahl überprüfter Datenelemente, Anzahl aller Patienten* |
| **Subkategorien** |
| Subkategorien zur Adjustierung des Indikators, optional |
| **Berechnung** |
| Beschreibung der Datensammlung und -aggregierung |
| **Interpretation** |
| Interpretation des Indikatorwertes, auch Vorschläge aus der Literatur<br>z.B. *Je höher die Rate, desto schlechter die Datenqualität. Meier/Schulze 2005 halten eine Rate < 0,5% für akzeptabel.* |
| **Einflussfaktoren** |
| Faktoren auf den Ebenen Struktur und Prozesse, die den Indikator beeinflussen können, sind für eine Schwachstellenanalyse hilfreich<br>z.B. *Einführung von Kategorien für not applicable/not done verringert die Anzahl echter fehlender Werte* |



## II.2.1 Indikatoren der Ebene Integrität

### II.2.1.1 Übereinstimmung mit Vorwert

**Beschreibung**

Übereinstimmung des Ergebnisses einer Messung oder Untersuchung mit dem Vorwert

**Begriffsdefinitionen**

*Vorwert:* Ein bezogen auf den aktuellen Wert zu einem früheren Zeitpunkt erhobener Wert desselben Datenelements.

**ID**

TMF-1001

**Typ**

Strukturqualität

**Quellen**

Gassmann et al. 1995, Naumann/Rolker 2000, O'Kelly 2004

**Bezug**

Die Berechnung kann für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für Datenelemente sinnvoll, die mehrfach erhoben werden (z.B. Laborparameter, Vitalparameter).

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator unterstützt die Sichtbarmachung von Datenfehlern. Werte, die sich in unplausibler Weise von ihren Vorwerten unterscheiden, sind ein deutlicher Hinweis auf das Vorliegen eines Datenfehlers.

Der Indikator ist mit dem Indikator „Reliabilität" verwandt.

Es muss hier unterschieden werden zwischen
- unveränderlichen Merkmalen, die sich überhaupt nicht (z.B. Geburtsdatum) oder nur im Rahmen der Messgenauigkeit (z.B. Körpergröße) ändern sollten, und
- veränderlichen Merkmalen, bei denen eine Abweichung vom Vorwert auftreten darf oder zu erwarten ist

**Zähler**

Anzahl der Werte mit absoluter Differenz zum Vorwert > akzeptable Abweichung

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

---

| Berechnung |
|---|
| 1) Für Datenelemente, die sich (im Rahmen der Messgenauigkeit) *nicht* ändern sollen: Festlegung der akzeptablen Abweichung pro Datenelement unter Berücksichtigung der Variabilität der Messungen (z.B. „0“ für das Geburtsdatum, „1 cm“ für die Körpergröße)<br>Für veränderliche Datenelemente: Festlegung der akzeptablen, unter den gegebenen Bedingungen des Registers plausiblen Abweichung pro Datenelement<br>2) Berechnung der absoluten Differenz zum Vorwert = Betrag (aktueller Wert – Vorwert)<br>3) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto mehr unplausible Werte (potenzielle Datenfehler) gibt es und umso schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Standardisierung von Untersuchungsverfahren und Analysemethoden; einmalige Erfassung unveränderlicher Merkmale |



## II.2.1.2 Konkordanz

**Beschreibung**

Konkordanz

**Begriffsdefinitionen**

*Konkordanz:* Die Übereinstimmung des Wertes eines Datenelements mit einer Referenzquelle (z.B. Patientenakte, CRF).

**ID**

TMF-1002

**Typ**

Strukturqualität

**Quellen**

Aronsky/Haug 2000, Brennan/Stead 2000, Cook et al. 2003

**Bezug**

Die Berechnung erfolgt für einen Datenbestand. Die Berechnung des Indikators ist nur für Datenelemente sinnvoll.

**Alternative Definition**

Konkordanz kann auch als Frage nach der Vollständigkeit der Datenerhebung betrachtet werden [Aronsky/Haug 2000], z.B. „sind alle in der Patientenakte vermerkten Diagnosen im Register gespeichert?“

**Anmerkungen**

Der Indikator überprüft nur die Übereinstimmung zweier Repräsentationen der Wirklichkeit und enthält keine Aussage über den Wahrheitsgehalt und die Angemessenheit der Darstellung [Brennan/Stead 2000]. Die Daten der Referenzquelle werden bei der Berechnung des Indikators als „wahr“ angenommen. Dies sollte bei der Wahl der Referenzquelle berücksichtigt werden.

Wenn die Referenzquelle mit der Quelle der Datenerhebung für das Register identisch ist, dann entspricht der Indikator dem Indikator „Übereinstimmung der Registerdaten mit den Originaldaten“.

**Zähler**

Anzahl der Werte, die nicht mit der Referenzquelle übereinstimmen

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

---

**Berechnung**

1) Wahl der Referenzquelle

2) Vergleich der Werte des Datenelements mit den entsprechenden Werten aus der Referenzquelle

3) Berechnung von Zähler, Nenner und Rate

| Interpretation |
| --- |
| Je höher die Rate, desto schlechter ist die Übereinstimmung mit der Referenzquelle (und damit die Datenqualität). Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Verfügbare Zeit und verfügbares Personal für die Datenerhebung; Qualifikation des Datenerhebungspersonals |



## II.2.1.3 Widerspruchsfreiheit

**Beschreibung**

Widerspruchsfreiheit

**Begriffsdefinitionen**

---

**ID**

TMF-1003

**Typ**

Strukturqualität

**Quellen**

Bobrowski et al. 1999, Gaus 2003, Jung/Winter 2000, Naumann/Rolker 2000

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

Möglichst widerspruchsfreie Daten sind wichtig für die Aussagekraft von Auswertungen.

**Zähler**

Anzahl der Werte, die mindestens eine Plausibilitätsregel verletzen

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

TMF-1035, TMF-1004, TMF-1005

**Berechnung**

1) Festlegung von einer oder mehreren Plausibilitätsregeln für jedes ausgewählte Datenelement [siehe auch Goertzen/Stausberg 2004].

Diese Regeln können z.B. folgende Form haben:

*B1 und/oder B2 und/oder B3 … → A*

mit *B1, B2, …, A*: Bedingungen, die Werte oder Wertelisten von Datenelementen beschreiben.

Beispiele:

*Schwangerschaft = „ja“ → Geschlecht = „weiblich“*

Alter > 50 oder Geschlecht = „männlich“ → Schwangerschaft = „na“

2) Anwendung der Plausibilitätsregeln auf die vorhandenen Werte

3) Berechnung von Zähler, Nenner und Rate

| Interpretation |
| --- |
| Je höher die Rate, desto mehr Widersprüche stecken in den Daten und desto schlechter ist damit die Datenqualität und die Nutzbarkeit der Daten. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Automatische Plausibilitätsprüfungen bei der Datenerfassung |



| II.2.1.4 Endlosüberlebende |
|---|
| **Beschreibung** |
| Anteil Endlosüberlebender an allen Patienten/Probanden |
| **Begriffsdefinitionen** |
| *Endlosüberlebender:* Patient oder Proband, für den zur Laufzeit eines Vorhabens niemals formal der Tod festgestellt wird (z.B. durch Todesbescheinigung oder Information vom Einwohnermeldeamt) |
| **ID** |
| TMF-1035 |
| **Typ** |
| Strukturqualität |
| **Quellen** |
| TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)<br>Hentschel/Katalinic A 2008 |
| **Bezug** |
| Die Berechnung kann für einen Datenbestand durchgeführt werden. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Dieser Indikator sagt etwas aus über die Vollständigkeit und Validität von Follow-ups. Wenn Patienten die zu erwartende Überlebenszeit für ihre Tumorerkrankung deutlich überschreiten, kann dies ein Hinweis auf fehlende Informationen zum Tod oder eine fehlerhafte Recherche sein. Dies gilt speziell für Tumore mit schlechter Prognose oder kurzer Überlebenszeit (z.B. Pankreaskarzinom). |
| **Zähler** |
| Anzahl Endlosüberlebender |
| **Nenner** |
| Anzahl Patienten/Probanden im Datenbestand insgesamt |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von TMF-1003. |
| **Berechnung** |
| 1) Identifizierung aller Endlosüberlebenden im Datenbestand<br>2) Bestimmung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Datenqualität. |
| **Einflussfaktoren** |
| Fehlende Informationen von den Meldestellen, unzureichendes Follow-up |

### II.2.1.5 Sichere Widersprüche/Fehler

**Beschreibung**

Durch Plausibilitätsprüfungen zwischen mehreren Merkmalen entdeckte Widersprüche und Fehler können nach ihrer Schwere (und der daraus resultierenden Reaktion) unterschieden werden. Dieser Indikator misst sichere Widersprüche und Fehler.

**Begriffsdefinitionen**

---

**ID**

TMF-1004

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Open European Nephrology Science Center (OpEN.SC), Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)

Hentschel/Katalinic 2008

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Sichere Widersprüche und Fehler müssen bei epidemiologischen Krebsregistern zwingend korrigiert werden. Unzureichende Korrekturen können in diesem Fall zur Zurückweisung einer Meldung führen (siehe Indikator TMF-1040: „Zurückgewiesene Meldungen“).

**Zähler**

Anzahl Plausibilitätsprüfungen, die einen sicheren Widerspruch oder Fehler zeigen

**Nenner**

Anzahl aller durchgeführten Plausibilitätsprüfungen

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1003.

**Berechnung**

1) Definition des Begriffes „sicherer Widerspruch/Fehler“ (d.h. Widersprüche oder Fehler, die eine große Bedeutung für das Vorhaben besitzen)

2) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität.

**Einflussfaktoren**

Motivation der Studienteilnehmer, Vollständigkeit und Korrektheit der primären Datenquellen



## II.2.1.6 Mögliche Widersprüche/Warnungen

**Beschreibung**

Durch Plausibilitätsprüfungen zwischen mehreren Merkmalen entdeckte Widersprüche und Fehler können nach ihrer Schwere (und der daraus resultierenden Reaktion) unterschieden werden. Dieser Indikator misst mögliche Widersprüche und Fehler.

**Begriffsdefinitionen**

---

**ID**

TMF-1005

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Open European Nephrology Science Center (OpEN.SC), Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)

Hentschel/Katalinic 2008

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Für epidemiologische Krebsregister liegen mögliche Widersprüche und Fehler dann vor, wenn Daten unplausibel aber prinzipiell möglich sind (z.B. Brustkrebs beim Mann).

**Zähler**

Anzahl Plausibilitätsprüfungen, die einen möglichen Widerspruch oder Fehler zeigen

**Nenner**

Anzahl aller durchgeführten Plausibilitätsprüfungen

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1003.

**Berechnung**

1) Definition des Begriffes „möglicher Widerspruch/Fehler“ (d.h. Widerspruch oder Fehler, der eine untergeordnete Bedeutung für das Vorhaben besitzt)

2) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität. Die Rate kann auch einen Hinweis auf unzutreffende Plausibilitätsprüfungen geben.

| Einflussfaktoren |
| --- |
| Motivation der Studienteilnehmer, Vollständigkeit und Korrektheit der primären Datenquellen, Umfang und Tiefe der Plausibilitätsprüfungen |



## II.2.1.7 Werteverteilung

**Beschreibung**

Werteverteilungen innerhalb und zwischen Zentren

**Begriffsdefinitionen**

---

**ID**

TMF-1006

**Typ**

Strukturqualität

**Quellen**

Hasford/Staib 1994, Schmidt et al. 1995

**Bezug**

Die Berechnung kann für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für *numerische* Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

---

**Zähler**

Anzahl auffälliger Werteverteilungen

**Nenner**

Anzahl überprüfter Datenelemente

**Subkategorien**

TMF-1007, TMF-1009, TMF-1010, TMF-1011, TMF-1052

**Berechnung**

1) Berechnung der Kenngrößen Minimum, Maximum, Quartile, Mittelwert, Median und Standardabweichung

2) Festlegung von Referenzwerten aus früheren Daten des Zentrums oder Daten der anderen Zentren

3) Vergleich der berechneten Kenngrößen mit den Referenzwerten

4) Berechnung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto mehr auffällige Werte stecken in den Daten und desto schlechter ist damit die Datenqualität. Jede festgestellte Auffälligkeit sollte zu einer vertieften Kontrolle der zugrunde liegenden Werte führen. Eine auffällige Verteilung kann ein Hinweis auf Probleme bei Untersuchungsverfahren und Analysemethoden (z.B. Laborparameter) oder auf Datenfälschung sein.

**Einflussfaktoren**

Standardisierung von Untersuchungsverfahren und Analysemethoden; unterschiedliche Patienten-/Personenkollektive in verschiedenen Zentren

### II.2.1.8 Bevorzugung bestimmter Endziffern

**Beschreibung**

Bevorzugung bestimmter Endziffern bei numerischen Datenelementen

**Begriffsdefinitionen**

*Endziffer:* Die am weitesten rechts stehende Ziffer eines Wertes. Dabei spielt die Stellung des Kommas keine Rolle (z.B. die Werte „10,43“ und „213“ besitzen die Endziffer „3“; der Wert „213,0“ besitzt die Endziffer „0“).

**ID**

TMF-1007

**Typ**

Strukturqualität

**Quellen**

Hasford/Staib 1994, O’Kelly 2004

**Bezug**

Die Berechnung erfolgt für einen Datenbestand. Die Berechnung des Indikators ist nur für *numerische* Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator basiert auf der Annahme, dass jede Ziffer aus dem Bereich von 0 bis 9 als Endziffer vorkommen kann und dass die Wahrscheinlichkeit für das Auftreten für alle Ziffern gleich groß ist (z.B. wenn das Körpergewicht mit einer Genauigkeit von 500 g in „kg“ gemessen wird, dann können als Endziffern nur „0“ und „5“ auftreten). Sind diese Bedingungen nicht erfüllt, sollte der Indikator nicht eingesetzt werden. Je kleiner die Anzahl überprüfter Werte, desto größer können auch bei korrekten Daten die Häufigkeitsunterschiede bei den Endziffern sein.

**Zähler**

Anzahl Datenelemente mit auffälliger Verteilung der Endziffern

**Nenner**

Anzahl überprüfter Datenelemente

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1006.

**Berechnung**

1) Berechnung der Häufigkeit des Auftretens der Endziffern 0 bis 9 pro ausgewähltem Datenelement

2) Untersuchung der gefundenen Häufigkeiten pro Datenelement auf auffällige (ungleichmäßige) Verteilung der Endziffern

3) Berechnung von Zähler, Nenner und Rate



| Interpretation |
| --- |
| Je höher die Rate, desto schlechter ist die Datenqualität. Auffällige Unterschiede können ein Hinweis auf Probleme bei Untersuchungsverfahren und Analysemethoden (z.B. Laborparameter) oder auf Datenfälschung sein und sollten eine vertiefte Kontrolle der zugrunde liegenden Werte nach sich ziehen. |
| **Einflussfaktoren** |
| Standardisierung von Untersuchungsverfahren und Analysemethoden |

## II.2.1.9 Werteverteilung der durch Untersucher erfassten Parameter

**Beschreibung**

Bei technischen Untersuchungen betrachtet SHIP die Qualität getrennt für das Gerät, den verantwortlichen Techniker (Observer/Untersucher) sowie die befundende Person (Reader). Dieser Indikator befasst sich mit der Qualität der Daten in Bezug auf den Untersucher.

**Begriffsdefinitionen**

---

**ID**

TMF-1009

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Study in Health in Pomerania (SHIP)

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator kann sowohl für einzelne Geräte als auch für einzelne technische Untersuchungen mit mehreren Geräten berechnet werden. Der Indikator ist nur dann sinnvoll, wenn als Zwischenschritt zwischen der Erzeugung von Messwerten und der Befundung eine manuelle Erfassung durch Untersucher stattfindet.

**Zähler**

Anzahl Untersucher mit auffälliger Werteverteilung

**Nenner**

Anzahl Untersucher

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1006.

**Berechnung**

1) Auswahl der technischen Untersuchung

2) Festlegung des Kriteriums für Auffälligkeit, z.B. Abweichung von einem Verteilungsparameter um einen bestimmten Betrag

3) Berechnung des Verteilungsparameters

4) Bestimmung von Zähler, Nenner und Rate



| Interpretation |
| --- |
| Je höher die Rate, desto schlechter ist die Datenqualität. Bei Einschluss mehrerer Geräte sind Verzerrungen durch feste Bindungen zwischen Untersucher und Geräten auszuschließen. Die Anzahl von Untersuchungen je Untersucher ist zu berücksichtigen. |
| **Einflussfaktoren** |
| Standardisierung der Erfassung von Werten, Schulung der Untersucher |

## II.2.1.10 Werteverteilung der durch Geräte erfassten Parameter

**Beschreibung**

Bei technischen Untersuchungen betrachtet SHIP die Qualität getrennt für das Gerät, den verantwortlichen Techniker (Observer) sowie die befundende Person (Reader). Dieser Indikator befasst sich mit der Qualität der Daten in Bezug auf die Geräte.

**Begriffsdefinitionen**

---

**ID**

TMF-1010

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Study in Health in Pomerania (SHIP)

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator wird für mehrere Geräte berechnet.

**Zähler**

Anzahl Geräte mit auffälliger Werteverteilung

**Nenner**

Anzahl Geräte

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1006.

**Berechnung**

1) Auswahl der Geräte

2) Festlegung des Kriteriums für Auffälligkeit, z.B. Abweichung von einem Verteilungsparameter um einen bestimmten Betrag

3) Berechnung des Verteilungsparameters

4) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität. Die Anzahl der Untersuchungen pro Gerät ist zu berücksichtigen.

**Einflussfaktoren**

Regelmäßige Wartung der Geräte, korrekte Einstellung und Handhabung der Geräte



## II.2.1.11 Werteverteilung von Befunden

**Beschreibung**

Bei technischen Untersuchungen betrachtet SHIP die Qualität getrennt für das Gerät, den verantwortlichen Techniker (Observer) sowie die befundende Person (Reader). Dieser Indikator befasst sich mit der Qualität der Daten in Bezug auf die befundende Person.

**Begriffsdefinitionen**

---

**ID**

TMF-1011

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Study in Health in Pomerania (SHIP)

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator kann sowohl für einzelne Geräte als auch für einzelne technische Untersuchungen mit mehreren Geräten berechnet werden.

**Zähler**

Anzahl Befunder mit auffälliger Werteverteilung

**Nenner**

Anzahl Befunder

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1006.

**Berechnung**

1) Auswahl der technischen Untersuchung

2) Festlegung des Kriteriums für Auffälligkeit, z.B. Abweichung von einem Verteilungsparameter um einen bestimmten Betrag

3) Berechnung des Verteilungsparameters

4) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität. Die Anzahl der Befunde je Befunder ist zu berücksichtigen.

**Einflussfaktoren**

Qualifikation der befundenden Personen, regelmäßige Zweitbefundung mit Feedback

### II.2.1.12 Werteverteilung von Parametern zwischen Zentren

**Beschreibung**

Messwerte können sich systematisch zwischen Zentren unterscheiden. Mögliche Unterschiede werden mit diesem Punkt abgebildet.

**Begriffsdefinitionen**

---

**ID**

TMF-1052

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Study in Health in Pomerania (SHIP)

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator wird für alle Zentren berechnet.

**Zähler**

Anzahl Zentren mit auffälliger Werteverteilung

**Nenner**

Anzahl Zentren

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1006.

**Berechnung**

1) Auswahl der Zentren

2) Festlegung des Kriteriums für Auffälligkeit, z.B. Abweichung von einem Verteilungsparameter um einen bestimmten Betrag

3) Berechnung des Verteilungsparameters

4) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität.

**Einflussfaktoren**

Schulungsvoraussetzungen, organisatorische Gegebenheiten, Gerätespezifika etc. in den Zentren.



## II.2.1.13 Untersuchungen am Wochenende

**Beschreibung**

Untersuchungen am Wochenende oder an gesetzlichen Feiertagen

**Begriffsdefinitionen**

---

**ID**

TMF-1008

**Typ**

Strukturqualität

**Quellen**

Hasford/Staib 1994

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für Datenelemente sinnvoll, die Datumsangaben von Untersuchungen repräsentieren.

**Alternative Definition**

---

**Anmerkungen**

Bei gesetzlichen Feiertagen ist zu prüfen, ob sie am Ort der Datenentstehung als normale Arbeitstage galten. Unter bestimmten Umständen sind auch Untersuchungen am Wochenende oder an gesetzlichen Feiertagen möglich (z.B. in der Notaufnahme eines Krankenhauses). Solche Fälle sind nicht als Datenfehler zu interpretieren.

**Zähler**

Anzahl der Datumsangaben, die auf ein Wochenende oder einen gesetzlichen Feiertag fallen.

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

---

**Berechnung**

1) Bestimmung der Datumsangaben aller Wochenenden und am Ort der Datenentstehung gültigen gesetzlichen Feiertage

2) Berechnung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. Jeder gefundene Fehler ist ein Hinweis auf einen Datenfehler oder Datenfälschung und sollte überprüft werden.

**Einflussfaktoren**

Automatische Plausibilitätsprüfungen bei der Datenerfassung

## II.2.1.14 Fehlende Module

**Beschreibung**

Der Datensatz eines Vorhabens kann in verschiedene Module unterteilt sein, z.B. nach Verfahren wie EKG oder nach sachlogischen Bezügen wie Soziodemographie. Module können auch über Gruppen von zusammenhängenden Variablen definiert sein. Dieser Indikator prüft die Vollständigkeit, in denen die Module vorliegen.

**Begriffsdefinitionen**

---

**ID**

TMF-1012

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Study in Health in Pomerania (SHIP)

**Bezug**

Die Berechnung kann für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator kann sowohl beschränkt auf ein Modul wie Soziodemographie als auch unter Einschluss verschiedener Module durchgeführt werden. Betrachtet werden nur komplette Module. Fehlende Module können z.B. durch fehlende Compliance von Studienteilnehmern in Bezug auf bestimmte Untersuchungen bedingt sein.

**Zähler**

Anzahl fehlender Module

**Nenner**

Anzahl überprüfter Module

**Subkategorien**

Der Indikator ist eine Subkategorie von „Fehlende Einträge“

**Berechnung**

1) Suche nach fehlenden Modulen

2) Berechnung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab.

**Einflussfaktoren**

Motivation der Studienteilnehmer, Belastung der Studienteilnehmer durch Untersuchungen, Geräteausfälle, Studiendesign



| II.2.1.15 Fehlende Werte bei Datenelementen |
|---|
| **Beschreibung** |
| Anteil fehlender Werte bei Datenelementen |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1013 |
| **Typ** |
| Strukturqualität |
| **Quellen** |
| TMF-Projekt V020-04: Ergänzung erfolgt aus Gründen der Systematik |
| **Bezug** |
| Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Dieser Indikator wurde eingeführt, um eine Berechnung der Anzahl fehlender Werte unter Einschluss von optionalen und mandatorischen Datenelementen zu ermöglichen. |
| **Zähler** |
| Anzahl fehlender Werte |
| **Nenner** |
| Anzahl überprüfter Werte |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von „Fehlende Einträge". Subkategorien dieses Indikators sind TMF-1014 und TMF-1015. |
| **Berechnung** |
| 1) Suche nach fehlenden Werten<br>2) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Einführung von separaten Kategorien für „not applicable", „not done". Nutzung eines EDC-Systems mit Erzwingung von Einträgen bei mandatorischen Datenelementen. |

## II.2.1.16 Fehlende Werte bei mandatorischen Datenelementen

**Beschreibung**

Anzahl fehlender Werte bei mandatorischen Datenelementen

**Begriffsdefinitionen**

---

**ID**

TMF-1014

**Typ**

Ergebnisqualität

**Quellen**

Aronsky/Haug 2000, Gaus 2003, Parkin/Muir 1992, Swart/Ihle 2005

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für *mandatorische* Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator ist wichtig für die korrekte Schätzung von Inzidenzen und Prävalenzen. Der Anteil der Fälle mit vollständigen Daten ist für spätere multivariate Analysen von großer Bedeutung.

Ein entsprechender Indikator ist auch für optionale Datenelemente vorhanden. Die separate Behandlung von mandatorischen und optionalen Datenelementen wird von Aronsky/Haug vorgeschlagen, da jeweils unterschiedliche Fehlerraten tolerabel sind [Aronsky/Haug 2000]. Gaus schlägt vor, die Untersuchung fehlender Werte auf mandatorische Datenelemente zu beschränken [Gaus 2003].

**Zähler**

Anzahl fehlender Werte bei mandatorischen Datenelementen

**Nenner**

Anzahl überprüfter Werte bei mandatorischen Datenelementen

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1013.

**Berechnung**

1) Suche nach fehlenden Werten bei mandatorischen Datenelementen

2) Berechnung von Zähler, Nenner und Rate

Alternative Berechnungsmöglichkeit: Anzahl fehlender Werte pro Patient/Person



| Interpretation |
|---|
| Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Einführung von separaten Kategorien für „not applicable“ und „not done“. |

## II.2.1.17 Fehlende Werte bei optionalen Datenelementen

**Beschreibung**

Anzahl fehlender Werte bei optionalen Datenelementen

**Begriffsdefinitionen**

---

**ID**

TMF-1015

**Typ**

Strukturqualität

**Quellen**

Aronsky/Haug 2000, Gaus 2003, Parkin/Muir 1992, Swart/Ihle 2005

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für *optionale* Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator ist wichtig für die korrekte Schätzung von Inzidenzen und Prävalenzen. Ein entsprechender Indikator ist auch für mandatorische Datenelemente vorhanden. Die separate Behandlung von mandatorischen und optionalen Datenelementen wird von Aronsky/Haug vorgeschlagen, da jeweils unterschiedliche Fehlerraten tolerabel sind [Aronsky/Haug 2000]. Gaus schlägt vor, die Untersuchung fehlender Werte auf mandatorische Datenelemente zu beschränken [Gaus 2003].

**Zähler**

Anzahl fehlender Werte bei optionalen Datenelementen

**Nenner**

Anzahl überprüfter Werte bei optionalen Datenelementen

**Subkategorien**

Der Indikator ist eine Subkategorie von TMF-1013.

**Berechnung**

1) Suche nach fehlenden Werten bei optionalen Datenelementen

2) Berechnung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab.

**Einflussfaktoren**

Einführung von separaten Kategorien für „not applicable“ und „not done“.



| II.2.1.18 Datenelemente mit Wert unbekannt o.ä. |
|---|
| **Beschreibung** |
| Anteil von Datenelementen mit Werten unbekannt o.ä. |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1016 |
| **Typ** |
| Strukturqualität |
| **Quellen** |
| TMF-Projekt V020-04: Study in Health in Pomerania (SHIP) |
| **Bezug** |
| Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Dieser Eintrag ist nur für mandatorische Datenelemente mit einer vorgegebenen Werteliste sinnvoll, bei denen ein Eintrag unbekannt o.ä. verfügbar ist. Dieser Indikator ist verwandt mit dem Indikator TMF-1025 „Datenelemente mit unspezifischen Werten“. |
| **Zähler** |
| Anzahl Werte unbekannt o.ä. |
| **Nenner** |
| Anzahl überprüfter Werte |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von „Fehlende Einträge“ |
| **Berechnung** |
| 1) Festlegung der relevanten Bezeichnungen („unbekannt“ o.ä.)<br>2) Suche nach den entsprechenden Einträgen<br>3) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Motivation des Erhebungspersonals, Art der Erhebung |

| II.2.1.19 Datenelemente mit bestehenden Einträgen bei allen Beobachtungseinheiten |
|---|
| **Beschreibung** |
| Anteil von Datenelementen mit bestehenden Einträgen bei allen Beobachtungseinheiten |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1017 |
| **Typ** |
| Strukturqualität |
| **Quellen** |
| TMF-Projekt V020-04: Open European Nephrology Science Center (OpEN.SC) |
| **Bezug** |
| Die Berechnung erfolgt für einen Datenbestand. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Die Berechnung ist sowohl für optionale als auch für mandatorische Datenelemente sinnvoll. Festzulegen ist der Umgang mit Einträgen wie „unbekannt“ o.ä. |
| **Zähler** |
| Anzahl von Datenelementen mit Einträgen bei allen Beobachtungseinheiten |
| **Nenner** |
| Anzahl von überprüften Datenelementen |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von „Fehlende Einträge“ |
| **Berechnung** |
| 1) Festlegung der einzuschließenden Datenelemente<br>2) Bestimmung für jedes Datenelement, ob Werte von allen Beobachtungseinheiten vorliegen<br>3) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je niedriger die Rate, desto schlechter die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Die Einflussfaktoren sind vielfältig, da dieser Indikator Aspekte anderer Kennzahlen bündelt. Es ist daher sinnvoll, diesen Indikator nicht getrennt, sondern zusammen mit anderen aus dem Bereich „Fehlende Einträge“ zu betrachten. |



## II.2.1.20 Ausreißer bei stetigen Datenelementen

**Beschreibung**

Ausreißer bei stetigen Datenelementen

**Begriffsdefinitionen**

*Stetiges Datenelement:* Ein Datenelement, das in einem bestimmten Bereich jeden Zwischenwert annehmen kann (z.B. Blutdruck).

*Ausreißer:* Wert eines Datenelements, der nicht den Erwartungen entspricht, d.h. außerhalb des Intervalls der erwarteten oder plausiblen Werte liegt.

**ID**

TMF-1018

**Typ**

Strukturqualität

**Quellen**

Gaus 2003, Swart/Ihle 2005

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für *stetige* Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator ist speziell für Laborparameter und Vitalparameter einsetzbar. Nicht jeder Ausreißer ist ein Datenfehler, z.B. können abhängig vom Schweregrad einer Erkrankung und dem Zustand eines Patienten sehr hohe oder sehr niedrige Laborwerte auftreten.

**Zähler**

Anzahl der Ausreißer

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

Je nach Datenelement möglich, z.B. Geschlecht bei Laborparametern

**Berechnung**

1) Festlegung des Intervalls erwarteter Werte pro Datenelement.

Ein bei der Darstellung von Box-Whisker-Plots benutztes Intervall ist der 1,5-fache Interquartilsabstand, d.h. der 1,5-fache Abstand zwischen dem 1. und 3. Quartil.

2) Suche nach Werten, die außerhalb dieses Intervalls liegen

3) Berechnung von Zähler, Nenner und Rate

| Interpretation |
| --- |
| Je höher die Rate, desto mehr Ausreißer und damit Datenfehler befinden sich in den Daten. Eine hohe Rate kann ein Hinweis auf Probleme mit den zugrunde liegenden Untersuchungs- oder Analysemethoden sein. |
| **Einflussfaktoren** |
| Standardisierung von Untersuchungsverfahren und Analysemethoden |



| II.2.1.21 Werte, die die Messbarkeitsgrenzen von Verfahren unter- oder überschreiten |
|---|
| **Beschreibung** |
| Anteil von Werten, die die Messbarkeitsgrenzen von Verfahren unter- oder überschreiten |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1019 |
| **Typ** |
| Strukturqualität |
| **Quellen** |
| TMF-Projekt V020-04: Study in Health in Pomerania (SHIP) |
| **Bezug** |
| Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Die Berechnung ist nur dann sinnvoll, wenn die Werte, die Messbarkeitsgrenzen unter- oder überschreiten, identifizierbar sind, z.B. durch eine bestimmte Kodierung. Die Auswertung kann verfahrensspezifisch erfolgen. Es kommen nur technische Verfahren in Betracht. Zudem müssen Messbarkeitsgrenzen in der Praxis relevant sein. |
| **Zähler** |
| Anzahl von Werten, die die Messbarkeitsgrenzen unter- oder überschreiten |
| **Nenner** |
| Anzahl von überprüften Werten |
| **Subkategorien** |
| --- |
| **Berechnung** |
| 1) Auswahl des Verfahrens<br>2) Suche nach Werten, die die Messbarkeitsgrenzen unter- oder überschreiten<br>3) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. Hohe Werte zeigen eine unangemessene Wahl des Verfahrens für das untersuchte Kollektiv an. |
| **Einflussfaktoren** |
| Verfahrensgerechte Durchführung der Untersuchungen, korrekte Indikationsstellung |

## II.2.1.22 Werte aus Standards

**Beschreibung**

Anteil von Werten, die mit Bezeichnungen aus kontrollierten Vokabularen übereinstimmen

**Begriffsdefinitionen**

---

**ID**

TMF-1020

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Open European Nephrology Science Center (OpEN.SC)

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

---

**Anmerkungen**

Die Berechnung ist nur sinnvoll für Freitextfelder oder Datenelemente, bei denen die Liste der erlaubten Werte bei der Erfassung ergänzt werden kann. Es ist sinnvoll, das zum Textvergleich eingesetzte Verfahren einer expliziten Qualitätsprüfung zu unterziehen.

**Zähler**

Anzahl von Werten mit Bezeichnungen aus kontrollierten Vokabularen.

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

---

**Berechnung**

1) Festlegung der zu berücksichtigenden kontrollierten Vokabulare

2) Durchführung einer Normalisierung der Einträge bzw. Festlegung eines Verfahrens zum Textvergleich

3) Suche nach Werten, die mit den Bezeichnungen aus kontrollierten Vokabularen übereinstimmen

4) Berechnung von Zähler, Nenner und Rate

**Interpretation**

Je niedriger die Rate, desto schlechter die Datenqualität. Da bei der Definition der Datenelemente auf die explizite Verwendung eines kontrollierten Vokabulars verzichtet wurde, ist eine Rate von 100% nicht zu erwarten. Die für diese Entscheidung relevanten Gründe müssen bei der Interpretation der Ergebnisse berücksichtigt werden.

**Einflussfaktoren**

Bedeutung und Bekanntheitsgrad von kontrollierten Vokabularen bei den betrachteten Datenelementen



## II.2.1.23 Unerlaubte Werte bei qualitativen Datenelementen

**Beschreibung**

Unerlaubte Werte bei qualitativen Datenelementen

**Begriffsdefinitionen**

*Qualitatives Datenelement:* Ein Datenelement, das nur eine endliche Zahl qualitativer Werte annehmen kann, d.h. Werte, die sich nicht quantifizieren lassen (z.B. nach ICD-10 kodierte Diagnose).

**ID**

TMF-1021

**Typ**

Strukturqualität

**Quellen**

Gaus 2003, Swart/Ihle 2005, Winter et al. 2003

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für *qualitative* Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator macht nur Sinn, wenn eine Erfassung von unerlaubten Werten technisch überhaupt möglich ist. Bei Datenelementen, die kodierte Werte enthalten, macht der Indikator nur eine Aussage über die *formale* Korrektheit der Kodierung (z.B. erlaubter Wertebereich = alle existierenden ICD-10-Kodes). Das (schwierige) Problem der *inhaltlichen* Korrektheit einer Kodierung wird von diesem Indikator nicht erfasst.

**Zähler**

Anzahl der unerlaubten Werte

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

Der Indikator ist eine Subkategorie von „Unerlaubte Werte". Je nach Datenelement sind weitere Subkategorien möglich, z.B. Geschlecht und Alter bei Datenelement „Schwangerschaft".

**Berechnung**

1) Festlegung der erlaubten Werte pro Datenelement

2) Suche nach unerlaubten Werten

3) Berechnung von Zähler, Nenner und Rate

**Interpretation**

Jeder unerlaubte Wert ist ein Datenfehler. Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab.

| Einflussfaktoren |
|---|
| Automatische Bereichsprüfungen bei der Datenerfassung |



## II.2.1.24 Unerlaubte Werte bei qualitativen Datenelementen zur Kodierung von Missings

**Beschreibung**

Anzahl unerlaubter Werte bei qualitativen Datenelementen zur Kodierung von Missings

**Begriffsdefinitionen**

---

**ID**

TMF-1022

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Study in Health in Pomerania (SHIP)

**Bezug**

Die Berechnung des Indikators ist nur für qualitative Datenelemente sinnvoll. Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator ergibt nur dann einen Sinn, wenn das Auftreten von Missings technisch überhaupt möglich und eine Kodierung vorgesehen ist.

**Zähler**

Anzahl der unerlaubten Werte

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

Der Indikator ist eine Subkategorie von „Unerlaubte Werte".

**Berechnung**

1) Festlegung der erlaubten Werte für Missings pro Datenelement

2) Suche nach unerlaubten Werten

3) Berechnung von Zähler, Nenner und Rate

**Interpretation**

Jeder unerlaubte Wert ist ein Datenfehler. Je höher die Rate, desto schlechter ist die Datenqualität.

**Einflussfaktoren**

Automatische Belegung von Missings mit definierten Kodes

## II.2.1.25 Unerlaubte Werte zur Kodierung von fehlenden Modulen

**Beschreibung**

Anzahl unerlaubter Werte zur Kodierung von fehlenden Modulen

**Begriffsdefinitionen**

---

**ID**

TMF-1023

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Study in Health in Pomerania (SHIP)

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator ist verwandt mit Indikator TMF-1012 „Fehlende Module".

**Zähler**

Anzahl der unerlaubten Werte

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

Der Indikator ist eine Subkategorie von „Unerlaubte Werte".

**Berechnung**

1) Festlegung der erlaubten Werte für fehlende Module
2) Suche nach unerlaubten Werten
3) Berechnung von Zähler, Nenner und Rate

**Interpretation**

Jeder unerlaubte Wert ist ein Datenfehler. Je höher die Rate, desto schlechter ist die Datenqualität.

**Einflussfaktoren**

Automatische Belegung von fehlenden Modulen mit definierten Kodes



## II.2.1.26 Unerlaubte Werte bei quantitativen Datenelementen zur Kodierung von Unter- oder Überschreitungen von Messbarkeitsgrenzen

| II.2.1.26 Unerlaubte Werte bei quantitativen Datenelementen zur Kodierung von Unter- oder Überschreitungen von Messbarkeitsgrenzen |
|---|
| **Beschreibung** |
| Unerlaubte Werte bei quantitativen Datenelementen zur Kodierung von Unter- oder Überschreitungen von Messbarkeitsgrenzen |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1024 |
| **Typ** |
| Strukturqualität |
| **Quellen** |
| TMF-Projekt V020-04: Study in Health in Pomerania (SHIP) |
| **Bezug** |
| Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Dieser Indikator ist verwandt mit Indikator TMF-1019 „Werte, die die Messbarkeitsgrenzen von Verfahren unter- oder überschreiten“. |
| **Zähler** |
| Anzahl der unerlaubten Werte |
| **Nenner** |
| Anzahl überprüfter Werte |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von „Unerlaubte Werte“. |
| **Berechnung** |
| 1) Festlegung der erlaubten Werte zur Kodierung von Unter- oder Überschreitungen von Messbarkeitsgrenzen<br>2) Suche nach unerlaubten Werten<br>3) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Jeder unerlaubte Wert ist ein Datenfehler. Je höher die Rate, desto schlechter ist die Datenqualität. |
| **Einflussfaktoren** |
| Automatische Belegung von Unter- oder Überschreitungen von Messbarkeitsgrenzen mit definierten Kodes |

## II.2.1.27 Datenelemente mit unspezifischen Werten

**Beschreibung**

Anzahl von Datenelementen mit unspezifischen Werten

**Begriffsdefinitionen**

---

**ID**

TMF-1025

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)

Hentschel/Katalinic 2008

**Bezug**

Die Berechnung kann für einen eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator beschreibt die Informationsdichte.

**Zähler**

Anzahl Datenelemente mit unspezifischer Angabe

**Nenner**

Anzahl überprüfter Datenelemente

**Subkategorien**

Bei epidemiologischen Krebsregistern wird dieser Indikator berechnet für

- Uterus-Malignome,
- Histologie,
- Lokalisation und
- Diagnosen (z.B. ICD-10-Kodes der Form „xxx.9“).

**Berechnung**

1) Festlegung der zu untersuchenden Datenelemente

2) Definition der Kategorie „unspezifisch“ für die zu untersuchenden Datenelemente

3) Bestimmung von Zähler, Nenner und Rate



| Interpretation |
| --- |
| Je höher die Rate, desto schlechter ist die Datenqualität. Auf der Ebene des einzelnen Datensatzes kann der Indikator u.U. ignoriert werden, da bei der Zusammenführung von Datensätzen für eine Beobachtungseinheit unspezifische Angaben in einem Datensatz durch spezifischere Angaben in einem anderen Datensatz überschrieben werden können. Bei Uterus-Malignomen sollte die Rate unter 5% liegen. |
| **Einflussfaktoren** |
| Genauigkeit der Diagnostik, Vollständigkeit und Korrektheit der primären Datenquellen |

## II.2.1.28 Beobachtungseinheiten mit unbekanntem Primärtumor

**Beschreibung**

Anteil der Beobachtungseinheiten mit unbekanntem Primärtumor

**Begriffsdefinitionen**

---

**ID**

TMF-1026

**Typ**

Strukturqualität

**Quellen**

TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)

Hentschel/Katalinic 2008

**Bezug**

Die Berechnung kann für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator hat eine enge Beziehung zum Indikator TMF-1025 „Datenelemente mit unspezifischen Werten“ und kann ggf. als weitere Präzisierung aufgefasst werden.

**Zähler**

Anzahl der Beobachtungseinheiten mit unbekanntem Primärtumor

**Nenner**

Anzahl aller überprüften Beobachtungseinheiten

**Subkategorien**

---

**Berechnung**

1) Identifizierung aller Beobachtungseinheiten mit unbekanntem Primärtumor

2) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität. Die Rate sollte unter 5% liegen.

**Einflussfaktoren**

Genauigkeit der Diagnostik, Vollständigkeit und Korrektheit der primären Datenquellen



| II.2.1.29 Nachweis bekannter Korrelationen |
|---|
| **Beschreibung** |
| Nachweis einer bereits bekannten Korrelation zwischen verschiedenen Datenelementen |
| **Begriffsdefinitionen** |
| *Korrelation:* Vorhandensein einer statistischen Beziehung zwischen zwei oder mehr quantitativen Datenelementen. |
| **ID** |
| TMF-1027 |
| **Typ** |
| Strukturqualität |
| **Quellen** |
| Hasford/Staib 1994 |
| **Bezug** |
| Die Berechnung erfolgt für einen Datenbestand. Die Berechnung des Indikators ist nur für *quantitative* Datenelemente sinnvoll. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| --- |
| **Zähler** |
| Anzahl erwarteter, aber nicht nachweisbarer Korrelationen |
| **Nenner** |
| Anzahl überprüfter Korrelationen |
| **Subkategorien** |
| --- |
| **Berechnung** |
| 1) Bestimmung, welche Datenelemente miteinander korreliert sein sollten<br>2) Festlegung des statistischen Verfahrens zur Überprüfung der Korrelation<br>3) Anwendung des ausgewählten Verfahrens auf die ausgewählten Datenelemente<br>4) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Datenqualität. Wenn der Nachweis einer erwarteten Korrelation fehlschlägt, ist dies ein Hinweis auf Datenfehler oder Datenfälschung (und damit schlechter Datenqualität). Diese Fälle sollten im Detail überprüft werden. |
| **Einflussfaktoren** |
| --- |

| II.2.1.30 Umfang der Metadaten bei Untersuchungen |
|---|
| **Beschreibung** |
| Anteil der von Untersuchungen übermittelten Metadaten |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1050 |
| **Typ** |
| Prozessqualität |
| **Quellen** |
| Metadaten: Daten, die andere Daten definieren und beschreiben (aus ISO/IEC 11179-3 Edition 3); hier verstanden als Parameter einer technischen Untersuchung. |
| **Bezug** |
| TMF-Projekt V020-04: Open European Nephrology Science Center (OpEN.SC) |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Dieser Indikator ist nur dann sinnvoll, wenn die Möglichkeit zur Auslassung von Parametern bei der Übermittlung von Ergebnissen technischer Untersuchungen besteht. |
| **Zähler** |
| Anzahl vorhandener Parameter |
| **Nenner** |
| Anzahl der zu erwartenden Parameter |
| **Subkategorien** |
| --- |
| **Berechnung** |
| 1) Festlegung der möglichen Parameter der interessierenden Untersuchung<br>2) Bestimmung der vorhandenen Parameter<br>3) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je niedriger die Rate, desto schlechter ist die Datenqualität. |
| **Einflussfaktoren** |
| Umfang von Schnittstellen, Einstellung von Geräten |



## II.2.2 Indikatoren der Ebene Organisation

### II.2.2.1 Aktualität der gespeicherten Daten

**Beschreibung**

Aktualität der gespeicherten Daten

**Begriffsdefinitionen**

---

**ID**

TMF-1028

**Typ**

Prozessqualität

**Quellen**

Abate et al. 1998, Bobrowski et al. 1999, Jung/Winter 2000, Katalinic 2005, Kuntoro et al. 1994, Naumann/Rolker 2000, Wang et al. 1993

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung ist nur für *veränderliche* Datenelemente sinnvoll.

**Alternative Definition**

„Aktualität" kann auch aufgefasst werden als Frage nach dem Alter von gespeicherten Daten [Naumann/Rolker 2000, Wang et al. 1993]. Dieser Ansatz ist nur sinnvoll für Daten, die sich über die Zeit verändern können und iterativ erhoben werden.

**Anmerkungen**

Der Indikator ist wichtig für die Aussagekraft von Auswertungen. Seine Bedeutung hängt von der Art der Daten und der geplanten Datennutzung ab.

Er ist nicht anwendbar bei unveränderlichen soziodemographischen (z.B. Geburtsdatum, Geburtsort, Geschlecht, Todesursache) und krankheitsbezogenen Daten (z.B. angeborene Krankheiten und Behinderungen, Diagnose bei chronischen und unheilbaren Krankheiten, Daten von früheren Krankenhausaufenthalten, Geburten und Fehlgeburten, Im- und Transplantationen, Unfälle, Amputationen). Bei veränderlichen krankheitsbezogenen Daten (z.B. Diagnostik, Dauer, Medikation und Therapie von Akuterkrankungen) ist er bedingt anwendbar, wenn diese Daten aus alten Dokumenten oder aus dem Gedächtnis der Patienten/Personen rekonstruiert werden müssen.

Der Indikator ist eventuell wichtig für Feasibilitystudien, wenn über veränderliche medizinische Daten über den Ein- oder Ausschluss von Patienten in klinische Studien entschieden werden soll (z.B. Dosisveränderungen im Therapieverlauf bei vorgegebenem Dosisintervall). Bei allen akuten (besonders bei ernsthaften) unerwünschten Ereignissen im Rahmen einer Medikations- oder Therapiestudie ist der Indikator als Prozessindikator wichtig. Er ist evident wichtig bei meldepflichtigen Erkrankungen.

**Zähler**

Anzahl der Werte mit Zeitdifferenz > akzeptable Zeitdifferenz

| Nenner |
| --- |
| Anzahl überprüfter Werte |
| **Subkategorien** |
| --- |
| **Berechnung** |
| 1) Festlegung der akzeptablen Zeitdifferenz/Verfallzeit der Daten (z.B. 10 Tage); diese Werte können für verschiedene Datenelemente oder Informationseinheiten unterschiedlich sein.<br>2) Berechnung der aktuellen Zeitdifferenz = Datum des Eingangs in das Register – Datum des Arztkontaktes, bei dem die Daten entstanden sind [Katalinic 2005]<br>3) Berechnung von Zähler, Nenner und Rate<br>Alternative Berechnungsmöglichkeit [Kuntoro et al. 1994]:<br>1) Festlegung einer akzeptablen Aktualitätsrate<br>2) Festlegung eines festen Zeitintervalls<br>3) Berechnung des Zählers = Anzahl von Meldungen in dem festgelegten Zeitintervall<br>3) Berechnung des Nenners = erwartete Anzahl von Meldungen in dem festgelegten Zeitintervall<br>Dieser Wert kann aus bekannten Prävalenzen/Inzidenzen abgeleitet werden.<br>4) Berechnung der Aktualitätsrate<br>5) Aktualitätsrate < akzeptable Aktualitätsrate bedeutet schlechte Datenqualität. Dies ähnelt der Definition des Indikators „Vollständigkeit“. |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Datenqualität. |
| **Einflussfaktoren** |
| Personelle Kapazitäten in den Zentren; Qualität der technischen Infrastruktur; Güte der Datenquelle |



## II.2.2.2 Dubletten im Datenbestand

**Beschreibung**

Anzahl Dubletten im Datenbestand

**Begriffsdefinitionen**

*Dublette:* Zwei oder mehr Informationseinheiten, die dasselbe Subjekt oder Objekt beschreiben und bei denen die Identität nachgewiesen werden kann oder – sofern der Nachweis nicht möglich ist – mit hinreichender Wahrscheinlichkeit angenommen werden kann.

*Hinreichende Wahrscheinlichkeit:* Eine unter Berücksichtigung von Datenqualitätsaspekten und beabsichtigter Datennutzung als ausreichend angesehene Wahrscheinlichkeit.

**ID**

TMF-1029

**Typ**

Prozessqualität

**Quellen**

Swart/Ihle 2005

**Bezug**

Die Berechnung kann für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung ist nur für abgrenzbare Informationseinheiten sinnvoll (z.B. Diagnosen).

**Alternative Definition**

---

**Anmerkungen**

Der Indikator ist wichtig für die korrekte Schätzung von Inzidenzen und Prävalenzen und mit dem Indikator „Anzahl Synonyme“ verwandt. Die Existenz einer Dublette bedeutet, dass Daten real mehrfach vorhanden sind. Bei Synonymen hingegen wird zwar ein Patient/eine Person unter verschiedenen Patientenidentifikatoren mehrfach geführt. Dies muss aber nicht bedeuten, dass seine/ihre Daten doppelt vorhanden sind. Beispielsweise könnte eine Dokumentationskraft bei einem neuen Patienten nach Beschaffung eines Patientenidentifikators die Datenerfassung begonnen haben und eine zweite Dokumentationskraft nach versehentlicher zweiter Beschaffung eines Patientenidentifikators die Datenerfassung unter diesem neuen Patientenindikator fortsetzen.

Die geschätzte Rate an Dubletten muss auf jeden Fall in die Bewertung von statistischen Kennzahlen (Signifikanzniveau) eingehen.

**Zähler**

Anzahl gefundener Dubletten

**Nenner**

Anzahl überprüfter Sets von Datenelementen

**Subkategorien**

---

| Berechnung |
|---|
| 1) Festlegung eines Sets von Datenelementen, deren Kombination die Erkennung von Dubletten erlaubt (z.B. Geschlecht, Geburtsdatum, Wohnort, Infektionsdatum, Prozeduren)<br>2) Festlegung der Nachweisgrenze für die Dublettensuche (z.B. übereinstimmende Werte bei allen Datenelementen des Sets; übereinstimmende Werte bei mindestens 6 von 8 Datenelementen des Sets)<br>3) Identifizierung von Dubletten durch Wertevergleich der Schlüsselelemente pro Informationseinheit<br>4) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Datenqualität. Swart/Ihle 2005 bezeichnen für die Sekundärdatenanalyse von Routinedaten des Gesundheitswesens eine Rate < 0,01% als akzeptabel. Diese Rate erscheint zu rigide und für große Register und Kohortenstudien nicht realitätskonform. Aus einer empirischen Stichprobe von 700 Patienten aus einem Zentrum des Kompetenznetzes Parkinson lässt sich eine realistische Rate von 2,5% schätzen (ermittelt durch G. Antony, IT-Koordination Kompetenznetz Parkinson). |
| **Einflussfaktoren** |
| Zusammenführung von Daten aus verschiedenen Informationsquellen im Register; Patienten/Personen, die mehrere meldende Zentren aufsuchen; spezielle Art der Erkrankung (z.B. Demenzen); Alter (Kinder); Qualität der technischen Infrastruktur (Möglichkeit, Zentrumswechsel und parallele Behandlung eines Patienten/Personen in mehreren Zentren zu melden); Anzahl der in einem Zentrum rekrutierten Patienten/Personen (die Ärzte können sich nicht alle Patienten/Personen merken); Fluktuation im Zentrum (der neue Arzt weiß nicht, ob sein Vorgänger den Patienten/die Person schon gesehen hat). |



## II.2.2.3 Rekrutierungsrate

**Beschreibung**

Rekrutierungsrate

**Begriffsdefinitionen**

*Rekrutierung:* Aufnahme von Patienten/Personen, die die Einschluss- und Ausschlusskriterien erfüllen, in ein Register.

*Einschlusskriterien:* Menge von Eigenschaften eines Patienten/einer Person, die *alle gleichzeitig* vorhanden sein müssen, damit er/sie in das Register aufgenommen werden kann (z.B. Vorliegen einer bestimmten Erkrankung).

*Ausschlusskriterien:* Menge von Eigenschaften eines Patienten/einer Person, die seine/ihre Aufnahme in das Register nicht erlauben. Das Vorhandensein von *mindestens einer* Eigenschaft verbietet die Aufnahme des Patienten/der Person in das Register (z.B. Alter < 18 Jahre).

**ID**

TMF-1030

**Typ**

Prozessqualität

**Quellen**

Goldberg et al. 1980, Hasford/Staib 1994, Kuntoro et al. 1994, Pogash et al. 2001, Pollock 1994, Svolba/Bauer 1999

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

---

**Anmerkungen**

Zur Erhöhung der Aussagekraft des Indikators müssten die Krankheitsschwere und die Schwierigkeit, die Patienteneinwilligung zu erlangen (z.B. für aufwendige, eventuell. auch schmerzhafte Untersuchungen), bei der Berechnung berücksichtigt werden.

**Zähler**

Anzahl rekrutierter Patienten/Personen im gewählten Zeitintervall

**Nenner**

Anzahl der „potenziell passenden“ (d.h. rekrutierbaren) Patienten/Personen im gewählten Zeitintervall

**Subkategorien**

TMF-1051

| Berechnung |
| --- |
| 1) Festlegung eines angemessenen Zeitintervalls (z.B. 3 Monate)<br>2) Berechnung von Zähler, Nenner und Rate<br>Alternative Berechnungsmöglichkeiten:<br>1) Festlegung der erwarteten Variabilität der Rekrutierungsrate pro gewähltem Zeitintervall<br>2) Bestimmung der aktuellen Rekrutierungsrate (Anzahl Patienten/Personen pro Zeitintervall) im gewählten Zeitintervall<br>3) Abweichung = Rekrutierungsrate im vorherigen Zeitintervall – Rekrutierungsrate im aktuellen Zeitintervall<br>Absolutbetrag der Abweichung > erwartete Variabilität ist ein Hinweis auf Probleme bei der Rekrutierung.<br>Ableitung einer Schätzung für die Vollständigkeit des Registers aus einer Stichprobe, und Extrapolation dieser Schätzung auf das gesamte Register [Goldberg et al. 1980].<br>Eine aus historischen Daten bekannte Inzidenz- oder Prävalenzrate wird dazu benutzt, für die Registerdaten eine erwartete Zahl von Fällen zu berechnen. Die Differenz zwischen der erwarteten Fallzahl und der tatsächlich im Register vorgefundenen Fallzahl ist ein Maß für die Vollständigkeit [Goldberg et al. 1980, Kuntoro et al. 1994]. |
| **Interpretation** |
| Je höher die Rate, desto besser ist die Datenqualität. Eine Rekrutierungsrate < Schwellenwert zeigt an, dass ein zu kleiner Anteil der rekrutierbaren Patienten/Personen tatsächlich in das Register aufgenommen wurde. Die Vollzähligkeit ist unzureichend. |
| **Einflussfaktoren** |
| Gezielte Ansprache und Motivation rekrutierbarer Patienten/Personen in den Zentren; personelle Kapazitäten in den Zentren; Bereitschaft der Patienten/Personen zur Teilnahme; Schweregrad der Erkrankung (Rekrutierungsrate sinkt mit zunehmender Schwere der Erkrankung und damit zunehmendem Aufwand der Erhebung) |



| II.2.2.4 DCO-Rate (Death Certificate Only) |
|---|
| **Beschreibung** |
| Anteil der Krebserkrankungen, für die nur eine Todesbescheinigung mit einem Hinweis auf die Krebserkrankung vorliegt. |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1051 |
| **Typ** |
| Prozessqualität |
| **Quellen** |
| TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)<br>Hentschel/Katalinic 2008 |
| **Bezug** |
| Die Berechnung kann für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Dieser Indikator beschreibt die Vollzähligkeit der Krebsregistrierung. Er ist der wichtigste Indikator für epidemiologische Krebsregister und ist dort zur Beurteilung eines Datenbestandes der entscheidende Indikator. Nach seinem Wert entscheiden Forschungsprojekte (z.B. Survivalanalysen auf regionaler Ebene) und internationale Datensammlungen (Globocan, Cancer Incidence in five continents), ob die Daten eines Krebsregisters verwendet werden oder nicht. Bei der Berechnung und Beurteilung anderer Indikatoren ist die DCO-Rate als Einflussgröße zu berücksichtigen. |
| **Zähler** |
| Anzahl der Krebserkrankungen, für die nur eine Todesbescheinigung mit einem Hinweis auf die Krebserkrankung vorliegt. |
| **Nenner** |
| Anzahl aller erfassten Krebserkrankungen |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von TMF-1030. Die einzelnen Tumorentitäten definieren Subkategorien dieses Indikators. |
| **Berechnung** |
| 1) Identifizierung der Krebserkrankungen, für die nur eine Todesbescheinigung mit einem Hinweis auf die Krebserkrankung vorliegt<br>2) Bestimmung von Zähler, Nenner und Rate |

| Interpretation |
| --- |
| Je höher die Rate, desto schlechter ist die Datenqualität. Die DCO-Rate sollte insgesamt unter 5% liegen. Für die einzelnen Tumorentitäten können die Referenzwerte davon abweichen. Bei jungen Krebsregistern wird die DCO-Rate vergleichsweise hoch sein, da auch noch Jahre nach Beginn der Registrierung Personen an oder mit Krebs sterben, die vor Beginn der Registrierung an Krebs erkrankt sind. |
| **Einflussfaktoren** |
| Bereitschaft der Meldestellen zur Meldung, fehlerhafte Zusammenführung von Meldungen |



| II.2.2.5 Verweigerungsrate von Untersuchungen |
|---|
| **Beschreibung** |
| Anteil von verweigerten Untersuchungen |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1031 |
| **Typ** |
| Prozessqualität |
| **Quellen** |
| TMF-Projekt V020-04: Study in Health in Pomerania (SHIP) |
| **Bezug** |
| Die Berechnung erfolgt für einen Datenbestand. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Der Indikator ist verwandt mit Indikator TMF-1012 „Fehlende Module". Voraussetzung zur Ermittlung des Indikators ist die Möglichkeit, eine Verweigerung von Untersuchungen explizit zu dokumentieren. Dieser Indikator kann sowohl für eine spezifische Untersuchung, z.B. MRT, als auch für verschiedene Untersuchungen zusammen bestimmt werden. |
| **Zähler** |
| Anzahl verweigerter Untersuchungen |
| **Nenner** |
| Anzahl überprüfter Untersuchungen |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von „Verweigerungsraten". |
| **Berechnung** |
| 1) Identifikation des Datenelements, in dem die Verweigerung erfasst wird, für jeden Untersuchungstyp<br>2) Bestimmung des Wertes, der eine Verweigerung anzeigt<br>3) Suche nach dem Auftreten dieses Wertes<br>4) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Studiendesign, Incentive für Studienteilnehmer, Belastung der Studienteilnehmer |

| II.2.2.6 Verweigerungsrate von Modulen |
|---|
| **Beschreibung** |
| Anteil von verweigerten Modulen |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1032 |
| **Typ** |
| Prozessqualität |
| **Quellen** |
| TMF-Projekt V020-04: Open European Nephrology Science Center (OpEN.SC), Study in Health in Pomerania (SHIP) |
| **Bezug** |
| Die Berechnung erfolgt für einen Datenbestand. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Der Indikator ist eng verwandt mit den Indikatoren TMF-1012 „Fehlende Module" und TMF-1031 „Verweigerungsrate von Untersuchungen". Voraussetzung zur Ermittlung des Indikators ist die Möglichkeit, eine Verweigerung von Modulen explizit zu dokumentieren. |
| **Zähler** |
| Anzahl verweigerter Module |
| **Nenner** |
| Anzahl überprüfter Module |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von „Verweigerungsraten". |
| **Berechnung** |
| 1) Identifikation des Datenelements, in dem die Verweigerung erfasst wird<br>2) Bestimmung des Wertes, der eine Verweigerung anzeigt<br>3) Suche nach dem Auftreten dieses Wertes<br>4) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Studiendesign, Incentive für Studienteilnehmer |



| II.2.2.7 Verweigerungsrate bei einzelnen Datenelementen |
|---|
| **Beschreibung** |
| Anteil von verweigerten Datenelementen |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1033 |
| **Typ** |
| Prozessqualität |
| **Quellen** |
| TMF-Projekt V020-04: Study in Health in Pomerania (SHIP) |
| **Bezug** |
| Die Berechnung erfolgt für einen Datenbestand. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Der Indikator ist komplementär zum Indikator TMF-1013 „Fehlende Werte bei Datenelementen" und dessen Subkategorien TMF-1014 und TMF-1015. Voraussetzung zur Ermittlung des Indikators ist die Möglichkeit, eine Verweigerung der Beantwortung von Datenelementen explizit zu dokumentieren. Es wird sich bei den Datenelementen regelhaft um Elemente einer Befragung handeln, kaum um einen isolierten Parameter aus einer technischen Untersuchung. |
| **Zähler** |
| Anzahl von Werten mit Verweigerung |
| **Nenner** |
| Anzahl überprüfter Werte |
| **Subkategorien** |
| Der Indikator ist eine Subkategorie von „Verweigerungsraten". |
| **Berechnung** |
| 1) Identifikation von Datenelementen, bei denen eine Verweigerung erfasst wird<br>2) Bestimmung des Wertes, der eine Verweigerung anzeigt<br>3) Suche nach dem Auftreten dieses Wertes<br>4) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Sensibilität des Datenelements, Art der Befragung, Schulung der Interviewer |

## II.2.2.8 Vorzeitig aus dem Register ausscheidende Beobachtungseinheiten

**Beschreibung**

Anzahl Patienten/Personen, die vorzeitig aus dem Register ausscheiden (Drop-out-Rate)

**Begriffsdefinitionen**

*Vorzeitiges Ausscheiden:* Aus der Sicht des Datenmanagements die vorzeitige Beendigung der Datensammlung für einen Patienten/eine Person im Register. Dies kann dadurch geschehen, dass er/sie seine/ihre Einwilligung zur Teilnahme am Register zurücknimmt oder der Kontakt zwischen Zentrum und Patient/Person abbricht (lost-to-follow-up).

**ID**

TMF-1034

**Typ**

Prozessqualität

**Quellen**

Expertise IMIBE

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

Patienten/Personen, die durch Tod aus dem Register ausscheiden, können aus dem Zähler ausgeschlossen werden.

**Anmerkungen**

Eine möglichst niedrige Drop-out-Rate ist wichtig für die Aussagekraft von Auswertungen. Der Indikator ist nur wichtig für Verlaufsregister und für Register, aus denen Patienten für klinische Studien rekrutiert werden sollen (Feasibilitystudien). Er kann aber auch ein qualitativer medizinischer Indikator sein (z.B. angeborene Erkrankung bzw. Behinderung – Todeszeitpunkt oder Therapiezeitpunkt – Todeszeitpunkt) oder etwas über die „Einwilligungsschwere“ und die Zufriedenheit der Patienten/Personen mit dem Register aussagen (z.B. Untersuchungen sind zu aufwendig, schmerzhaft oder unangenehm, „es passiert ja nichts“). Bei Alterskrankheiten ist mit einer schwer zu schätzenden Dunkelziffer bei Todesfällen zu rechnen.

**Zähler**

Anzahl Patienten/Personen, die innerhalb des gewählten Zeitintervalls vorzeitig aus dem Register ausscheiden

**Nenner**

Anzahl im gewählten Zeitintervall neu in das Register aufgenommener Patienten/Personen

**Subkategorien**

---

**Berechnung**

1) Festlegung eines angemessenen Zeitintervalls (z.B. 3 Monate)

2) Berechnung von Zähler, Nenner und Rate

| Interpretation |
|---|
| Drop-out-Rate > Schwellenwert ist ein Hinweis auf Probleme im Zusammenhang mit Patientenansprache und -kontakt und kann zusätzlich ein Hinweis auf manipulierte Daten sein. Da (zu) viele Daten/Krankheitsverläufe nur unvollständig erfasst wurden, wird die Qualität und Aussagekraft der Daten dadurch insgesamt gemindert.<br>Im Kompetenznetz Parkinson wurde mit Stand 15.11.2005 eine Drop-out-Rate von 0,013% (62 von 4.801 Patienten) empirisch ermittelt (durch G. Antony, IT-Koordination Kompetenznetz Parkinson). Von den 62 Patienten waren 42 verstorben, die restlichen 20 Patienten zogen die Einwilligung für die weitere Teilnahme am Register zurück. |
| **Einflussfaktoren** |
| Maßnahmen zur Aufrechterhaltung und Verbesserung des Kontaktes zwischen Zentrum und Patienten/Personen; Information der Patienten/Personen über Sinn und Zweck des Registers; Personalkapazität in den Zentren (Stetigkeit) |

## II.2.2.9 Synonyme

**Beschreibung**

Anzahl Synonyme

**Begriffsdefinitionen**

*Synonym:* Lexikalisches Zeichen, das die gleiche Bedeutung wie ein anderes lexikalisches Zeichen hat. Im Zusammenhang mit Registern werden Patientenidentifikatoren als Synonyme bezeichnet, wenn sie sich lexikalisch unterscheiden, aber demselben Patienten/derselben Person zugewiesen wurden.

**ID**

TMF-1036

**Typ**

Prozessqualität

**Quellen**

Parkin/Muir 1992, Winter et al. 2003

**Bezug**

Die Berechnung erfolgt für einen Datenbestand. Falls möglich, sollten alle Patienten/Personen in die Suche nach Synonymen einbezogen werden.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator ist wichtig für die korrekte Schätzung von Inzidenzen und Prävalenzen und mit dem Indikator „Anzahl Dubletten im Datenbestand" verwandt. Nicht erkannte Synonyme führen dazu, dass Patienten/Personen mehrfach gezählt werden. Die Beziehungen zwischen Synonymen sollten dokumentiert werden [Winter et al. 2003].

**Zähler**

Anzahl Synonyme

**Nenner**

Anzahl überprüfter Patienten/Personen

**Subkategorien**

---

**Berechnung**

1) Festlegung der Datenelemente (Schlüsselelemente), deren Kombination die Erkennung von Patienten/Personen erlaubt, die mehrfach im Register gespeichert sind. Dazu können z.B. die Datenelemente gehören, die zur Erzeugung der Patientenidentifikatoren benutzt werden. Alternativ können Verfahren des Record Linkage eingesetzt werden.

2) Identifizierung der Synonyme durch Wertevergleich der Schlüsselelemente pro Patient/Person

3) Berechnung von Zähler, Nenner und Rate



| Interpretation |
|---|
| Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Algorithmus zur Erzeugung der Patientenidentifikatoren; Qualität der technischen Infrastruktur (Möglichkeit, Zentrumswechsel und parallele Behandlung eines Patienten in mehreren Zentren zu melden) |

### II.2.2.10 Homonyme

**Beschreibung**

Anzahl Homonyme

**Begriffsdefinitionen**

*Homonym:* Mit einem Wort sind zwei unterschiedliche Definitionen verbunden. Im Zusammenhang mit Registern wird ein Patientenidentifikator als Homonym bezeichnet, wenn er mehr als einem Patienten/einer Person zugewiesen wurde.

**ID**

TMF-1037

**Typ**

Prozessqualität

**Quellen**

Parkin/Muir 1992, Winter et al. 2003

**Bezug**

Die Berechnung erfolgt für einen Datenbestand. Falls möglich, sollten alle Patienten/Personen in die Suche nach Homonymen einbezogen werden.

**Alternative Definition**

---

**Anmerkungen**

Der Indikator ist wichtig für die korrekte Schätzung von Inzidenzen und Prävalenzen. Nicht erkannte Homonyme führen dazu, dass Daten verschiedener Patienten/Personen fälschlicherweise zusammengeführt werden. Dies kann zusätzlich zu vermeintlich falschen oder widersprüchlichen Daten führen und damit die Einschätzung der Datenqualität negativ beeinflussen. Die Entdeckung von Homonymen ist schwierig und unter Umständen nur eingeschränkt oder gar nicht möglich.

**Zähler**

Anzahl Homonyme

**Nenner**

Anzahl überprüfter Patienten/Personen

**Subkategorien**

---

**Berechnung**

1) Festlegung der Kriterien, anhand derer Homonyme erkannt werden können (z.B. die zur Erzeugung der Patientenidentifikatoren benutzten Datenelemente, doppelte Untersuchungen, Therapiebeginn vor Diagnosestellung, Therapie passt nicht zur Diagnose)

2) Identifizierung der Homonyme durch Überprüfung der festgelegten Kriterien

3) Berechnung von Zähler, Nenner und Rate



| Interpretation |
| --- |
| Je höher die Rate, desto schlechter ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Algorithmus zur Erzeugung der Patientenidentifikatoren |

## II.2.2.11 Einzelmeldungen pro Beobachtungseinheit

**Beschreibung**

Zu einem Tumorfall können mehrere Meldungen vorliegen. Der Indikator prüft, wie häufig zu einem Tumorfall nur eine einzige Meldung vorliegt.

**Begriffsdefinitionen**

---

**ID**

TMF-1038

**Typ**

Prozessqualität

**Quellen**

TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)

Hentschel/Katalinic 2008

**Bezug**

Die Berechnung kann für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Zu einem Tumorfall können mehrere Meldungen (z.B. vom niedergelassenen Arzt, Klinikarzt und Pathologen) vorliegen. Wenn nur eine einzelne Meldung zu einem Fall vorliegt, kann dies ein Hinweis auf schlechte Datenqualität sein, hervorgerufen durch fehlende oder unvollständige Meldungen. Dies gilt speziell für solitäre Meldungen von Pathologen, die hier mit berücksichtigt werden, aber auch durch einen separaten Qualitätsindikator (TMF-1039) zusätzlich untersucht werden können.

**Zähler**

Anzahl Beobachtungseinheiten mit nur einer Meldung

**Nenner**

Anzahl überprüfter Beobachtungseinheiten

**Subkategorien**

---

**Berechnung**

1) Bestimmung der Anzahl Meldungen für jede Beobachtungseinheit

2) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität.

**Einflussfaktoren**

Vollzähligkeit der Meldungen, Motivation der Meldestellen zur Meldung



## II.2.2.12 Solitäre Meldungen von Pathologen

| II.2.2.12 Solitäre Meldungen von Pathologen |
| --- |
| **Beschreibung** |
| Zu einem Tumorfall können mehrere Meldungen vorliegen. Der Indikator prüft, wie häufig zu einem Tumorfall nur eine solitäre Meldung eines Pathologen vorliegt. |
| **Begriffsdefinitionen** |
| --- |
| **ID** |
| TMF-1039 |
| **Typ** |
| Prozessqualität |
| **Quellen** |
| TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)<br>Hentschel/Katalinic 2008 |
| **Bezug** |
| Die Berechnung kann für einen Datenbestand durchgeführt werden. |
| **Alternative Definition** |
| --- |
| **Anmerkungen** |
| Zu einem Tumorfall können mehrere Meldungen (z.B. vom niedergelassenen Arzt, Klinikarzt und Pathologen) vorliegen. Wenn nur eine einzelne Pathologen-Meldung zu einem Fall vorliegt, bedeutet dies häufig eine schlechte Datenqualität, da in einer Pathologen-Meldung wichtige Informationen (z.B. Datum der Erstdiagnose), die dem Pathologen nicht bekannt sind, fehlen. Ein separater Qualitätsindikator (TMF-1038) überprüft solitäre Meldungen unabhängig vom Typ der Meldestelle. |
| **Zähler** |
| Anzahl Beobachtungseinheiten mit solitärer Pathologen-Meldung |
| **Nenner** |
| Anzahl überprüfter Beobachtungseinheiten |
| **Subkategorien** |
| --- |
| **Berechnung** |
| 1) Bestimmung der Anzahl Beobachtungseinheiten mit einer solitären Pathologen-Meldung<br>2) Bestimmung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Datenqualität. |
| **Einflussfaktoren** |
| Vollzähligkeit der Meldungen, Motivation der Meldestellen zur Meldung |

## II.2.2.13 Zurückgewiesene Meldungen

**Beschreibung**

Meldungen, die aufgrund fehlerhafter Angaben nicht verarbeitet werden können, werden zurückgewiesen. Der Indikator misst den Anteil solcher Meldungen.

**Begriffsdefinitionen**

---

**ID**

TMF-1040

**Typ**

Prozessqualität

**Quellen**

TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)

Hentschel/Katalinic 2008

**Bezug**

Die Berechnung kann für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Meldungen, die aufgrund fehlerhafter Angaben nicht verarbeitet werden können, werden zurückgewiesen. Die meldende Stelle hat dann die Möglichkeit, die Meldung mit korrigierten Daten erneut vorzunehmen.

**Zähler**

Anzahl zurückgewiesener Meldungen

**Nenner**

Anzahl überprüfter Meldungen

**Subkategorien**

---

**Berechnung**

1) Identifizierung zurückgewiesener Meldungen

2) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität.

**Einflussfaktoren**

Vollständigkeit und Korrektheit der primären Datenquellen, Motivation der Melder



## II.2.2.14 Datenquellen pro Beobachtungseinheit

**Beschreibung**

Zu einem Tumorfall können Meldungen aus einer oder mehreren Datenquellen vorliegen. Der Indikator prüft, wie häufig zu einem Tumorfall nur eine einzige Datenquelle existiert.

**Begriffsdefinitionen**

---

**ID**

TMF-1041

**Typ**

Prozessqualität

**Quellen**

TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)

Hentschel/Katalinic 2008

**Bezug**

Die Berechnung kann für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Zu einem Tumorfall können Meldungen aus mehreren Datenquellen (z.B. vom niedergelassenen Arzt, Klinikarzt und Pathologen) vorliegen. Wenn nur eine einzelne Datenquelle zu einem Fall zur Verfügung steht, fehlen möglicherweise wichtige Daten, da sie in dieser Datenquelle nicht erfasst wurden. Die Todesbescheinigung darf nicht mit in die Berechnung eingeschlossen werden.

**Zähler**

Anzahl Beobachtungseinheiten mit nur einer Datenquelle

**Nenner**

Anzahl überprüfter Beobachtungseinheiten

**Subkategorien**

---

**Berechnung**

1) Bestimmung der Anzahl Datenquellen für jede Beobachtungseinheit

2) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität.

**Einflussfaktoren**

Motivation der Meldestellen zur Meldung, Zeitpunkt der Berechnung

## II.2.2.15 Beobachtungseinheiten mit Follow-up

**Beschreibung**

Anteil Beobachtungseinheiten mit Follow-up

**Begriffsdefinitionen**

---

**ID**

TMF-1042

**Typ**

Prozessqualität

**Quellen**

TMF-Projekt V020-04: Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (GEKID)

Hentschel/Katalinic 2008

**Bezug**

Die Berechnung kann für einen Datenbestand durchgeführt werden.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator sagt etwas aus über die Vollzähligkeit von Follow-ups. Follow-ups sind wichtig für Verlaufsuntersuchungen (z.B. Bestimmung der Überlebenszeit nach Erstdiagnose eines Tumors). Voraussetzung ist die Kenntnis über die durchzuführenden Follow-ups.

**Zähler**

Anzahl Beobachtungseinheiten mit Follow-up

**Nenner**

Anzahl überprüfter Beobachtungseinheiten

**Subkategorien**

---

**Berechnung**

1) Identifizierung der Beobachtungseinheiten, für die mindestens ein Follow-up vorhanden ist

2) Bestimmung von Zähler, Nenner und Rate

**Interpretation**

Je geringer die Rate, desto schlechter ist die Datenqualität.

**Einflussfaktoren**

Bereitschaft der Meldestellen zur Meldung, Mobilität der Patienten



## II.2.3 Indikatoren der Ebene Richtigkeit

### II.2.3.1 Genauigkeit der Registerdaten

**Beschreibung**

Genauigkeit der Registerdaten

**Begriffsdefinitionen**

*Genauigkeit:* Das Ausmaß, in dem die Registerdaten der Wahrheit entsprechen [Arts et al. 2002a].

**ID**

TMF-1043

**Typ**

Ergebnisqualität

**Quellen**

Arts et al. 2002a, Barrie/Marsh 1992, Brennan/Stead 2000, Hassey et al. 2001, Hüsers 2004, Jung/Winter 2000, Kühn-Stoffers 2004, Lindquist 2004, Massey/Hoffman 1989, Prins et al. 2000

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für Datenelemente sinnvoll.

**Alternative Definition**

Genauigkeit kann auch definiert werden als Kombination aus den Indikatoren „Korrektheit" und „Vollständigkeit" [Brennan/Stead 2000, Hüsers 2004] oder als „Positiver Prädiktiver Wert" [Hassey et al. 2001]. Letzteres ist definiert als der Anteil der Patienten/Personen, die ein im Register gespeichertes Merkmal tatsächlich aufweisen, unter der Gesamtheit aller Patienten/Personen, für die dieses Merkmal im Register gespeichert ist.

**Anmerkungen**

Die „Wahrheit" sollte aus einer unabhängigen Datenquelle stammen, die nicht für die Erhebung der Registerdaten benutzt wurde. Der Vergleich der Registerdaten mit den Datenquellen, aus denen die Registerdaten abgeleitet wurden, wird durch den Indikator „Übereinstimmung der Registerdaten mit den Originaldaten" erfasst.

**Zähler**

Anzahl der im Register vorhandenen und korrekten Werte

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

–

**Berechnung**

1) Abstraktion der „Wahrheit" aus der unabhängigen Datenquelle

2) Überprüfung, ob diese Werte im Register vorhanden sind

3) Berechnung von Zähler, Nenner und Rate

| Interpretation |
|---|
| Je höher die Rate, desto besser ist die Datenqualität. Prins schlägt folgende Klassifikation vor [Prins et al. 2000]:<br>> 95%: gut<br>zwischen 80% und 95%: mäßig<br>< 80%: schlecht |
| **Einflussfaktoren** |
| Verfügbare Zeit und verfügbares Personal für die Datenerhebung; Qualifikation des Datenerhebungspersonals; Qualität (Wahrheitsgehalt) der für die Erhebung der Registerdaten benutzten Datenquellen |



## II.2.3.2 Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Datenelemente

**Beschreibung**

Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Datenelemente

**Begriffsdefinitionen**

*Originaldaten:* Die Daten, aus denen die Registerdaten extrahiert wurden. Die Originaldaten können in Papierform oder als elektronische Dokumente existieren.

**ID**

TMF-1044

**Typ**

Ergebnisqualität

**Quellen**

Arts et al. 2001, Arts et al. 2002a, Califf et al. 1997, Cnattingius et al. 1990, Gissler et al. 1996, Håkansson et al. 2001, Jensen et al. 2002, Khosla et al. 2000, Lin et al. 2004, Lu et al. 1995, Mullooly 1990, Maruszewksi et al. 2005, Nielsen et al. 1996, Ose et al. 2004, O'Sullivan et al. 1996, Teperi 1993, Pogash et al. 2001, Schmidt et al. 1995, Vestberg et al. 1997, Weiss et al. 1993, Weiss 1998

**Bezug**

Die Berechnung erfolgt für einen Datenbestand. Die Berechnung des Indikators ist nur für Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator bildet das Ergebnis einer Source Data Verification (Originaldatenabgleich) ab und wird auch für die Fallzahlplanung der Source Data Verification benötigt. Der Indikator überprüft nur die Übereinstimmung der Registerdaten mit den Originaldaten und enthält keine Aussage über den Wahrheitsgehalt der Daten. Die Originaldaten werden bei der Berechnung des Indikators als „wahr" angenommen. Dann können bei Nichtübereinstimmung die Registerdaten nur noch Datenfehler oder fehlende Werte sein.

Der Indikator ist verwandt mit den Indikatoren „Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Beobachtungseinheiten", „Konkordanz" und „Genauigkeit der Registerdaten".

**Zähler**

Anzahl der Werte, die nicht mit den Originaldaten übereinstimmen

**Nenner**

Anzahl überprüfter Werte

**Subkategorien**

---

| Berechnung |
| --- |
| 1) Vergleich der Werte des Datenelements mit den entsprechenden Werten aus den Originaldaten<br>2) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Übereinstimmung mit den Originaldaten (und damit die Datenqualität). Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. Eine hohe Rate kann ein Hinweis auf Probleme bei der Erhebung oder Erfassung der Registerdaten sein. |
| **Einflussfaktoren** |
| Verfügbare Zeit und verfügbares Personal für die Datenerhebung; Qualifikation des Datenerhebungspersonals; von der Ersterfassung unabhängige Zweiterfassung der Registerdaten |



## II.2.3.3 Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Beobachtungseinheiten

**Beschreibung**

Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Beobachtungseinheiten

**Begriffsdefinitionen**

*Originaldaten:* Die Daten, aus denen die Registerdaten extrahiert wurden. Die Originaldaten können in Papierform oder als elektronische Dokumente existieren.

**ID**

TMF-1045

**Typ**

Ergebnisqualität

**Quellen**

Arts et al. 2001, Arts et al. 2002a, Califf et al. 1997, Cnattingius et al. 1990, Gissler et al. 1996, Håkansson et al. 2001, Jensen et al. 2002, Khosla et al. 2000, Lin et al. 2004, Lu et al. 1995, Mullooly 1990, Maruszewksi et al. 2005, Nielsen et al. 1996, Ose et al. 2004, O'Sullivan et al. 1996, Teperi 1993, Pogash et al. 2001, Schmidt et al. 1995, Vestberg et al. 1997, Weiss et al. 1993, Weiss 1998

**Bezug**

Die Berechnung erfolgt für einen Datenbestand. Die Berechnung des Indikators ist nur für Datenelemente sinnvoll.

**Alternative Definition**

---

**Anmerkungen**

Dieser Indikator bildet das Ergebnis einer Source Data Verification (Originaldatenabgleich) ab und wird auch für die Fallzahlplanung der Source Data Verification benötigt. Der Indikator überprüft nur die Übereinstimmung der Registerdaten mit den Originaldaten und enthält keine Aussage über den Wahrheitsgehalt der Daten. Die Originaldaten werden bei der Berechnung des Indikators als „wahr" angenommen. Dann können bei Nichtübereinstimmung die Registerdaten nur noch Datenfehler oder fehlende Werte sein.

Der Indikator ist verwandt mit den Indikatoren „Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Datenelemente", „Konkordanz" und „Genauigkeit der Registerdaten".

**Zähler**

Anzahl Beobachtungseinheiten mit mindestens einem Fehler

**Nenner**

Anzahl überprüfter Beobachtungseinheiten

**Subkategorien**

---

| Berechnung |
|---|
| 1) Festlegung der Beobachtungseinheit (z.B. Patient, Meldung, Meldebogen)<br>2) Bestimmung der Anzahl Fehler für jede Beobachtungseinheit<br>3) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto schlechter ist die Übereinstimmung mit den Originaldaten (und damit die Datenqualität). Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. Eine hohe Rate kann ein Hinweis auf Probleme bei der Erhebung oder Erfassung der Registerdaten sein. |
| **Einflussfaktoren** |
| Verfügbare Zeit und verfügbares Personal für die Datenerhebung; Qualifikation des Datenerhebungspersonals; von der Ersterfassung unabhängige Zweiterfassung der Registerdaten |



## II.2.3.4 Vollständigkeit der Registerdaten

**Beschreibung**

Vollständigkeit der Registerdaten

**Begriffsdefinitionen**

*Vollständigkeit:* Hier das Ausmaß, in dem die notwendigen Daten, die in das Register eingehen könnten, auch tatsächlich eingegangen sind [Arts et al. 2002a].

**ID**

TMF-1046

**Typ**

Ergebnisqualität

**Quellen**

Arts et al. 2002a, Barrie/Marsh 1992, Bobrowski et al. 1999, Goldberg et al. 1980, Hassey et al. 2001, Hogan/Wagner 1997, Jensen et al. 2002, Jung/Winter 2000, Katalinic 2005, Kuntoro et al. 1994, Lindquist 2004, Logan et al. 2001, Naumann/Rolker 2000, Nielsen et al. 1996, Parkin/Muir 1992, Teppo et al. 1994, Topp et al. 1997, Vestberg et al. 1997

**Bezug**

Die Berechnung kann für einen einzelnen Datensatz, für eine einzelne Beobachtungseinheit und für einen Datenbestand durchgeführt werden. Die Berechnung des Indikators ist nur für Datenelemente sinnvoll.

**Alternative Definition**

Vollständigkeit kann auch definiert werden als Anteil Patienten/Personen mit vollständigen Daten oder als Sensitivität [Hassey et al. 2001, Nielsen et al. 1996], d.h. als der Anteil der Patienten/Personen, die ein im Register gespeichertes Merkmal tatsächlich aufweisen, unter der Gesamtheit aller Patienten/Personen, die dieses Merkmal in der Realität besitzen.

**Anmerkungen**

Dieser Indikator ist wichtig für die korrekte Schätzung von Inzidenzen und Prävalenzen. Er ist mit den Indikatoren „Konkordanz“ und „Genauigkeit der Registerdaten“ verwandt. Die Vollständigkeit der Rekrutierung wird durch den Indikator „Rekrutierungsrate“ erfasst. Als Quellen für die Festlegung der Daten, die in das Register eingehen könnten, können andere Register, Todesbescheinigungen, Abrechnungsdaten der Krankenkassen etc. herangezogen werden.

**Zähler**

Anzahl der im Register vorhandenen Werte [Naumann/Rolker 2000]

**Nenner**

Anzahl der im Register vorhandenen Werte + Anzahl der im Register fehlenden Werte (= Anzahl der Werte, die in das Register eingehen könnten) [Naumann/Rolker 2000]

**Subkategorien**

–

| Berechnung |
| --- |
| 1) Identifizierung der Daten bzw. Fälle, die in das Register eingehen sollten<br>2) Prüfung, ob diese Daten bzw. Fälle im Register vorhanden sind<br>3) Berechnung von Zähler, Nenner und Rate |
| **Interpretation** |
| Je höher die Rate, desto vollständiger sind die Daten und desto besser ist die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Verfügbare Zeit und verfügbares Personal für die Datenerhebung; Qualifikation des Datenerhebungspersonals; Vollständigkeit der den Registerdaten zugrunde liegenden Originaldokumente |



## II.2.3.5 Übereinstimmung mit Verfahrensregeln

**Beschreibung**

Übereinstimmung mit dem Studienplan

**Begriffsdefinitionen**

*Studienplan:* Ein oder mehrere Dokumente, die alle wesentlichen Aspekte eines Registers oder einer Kohortenstudie beschreiben (z.B. Einschluss- und Ausschlusskriterien für Fälle, Meldezeiträume, Zeitintervalle für Untersuchungen, Untersuchungs- und Analysemethoden)

**ID**

TMF-1047

**Typ**

Ergebnisqualität

**Quellen**

DAE 2004, Gaus 2003, Whitney et al. 1998

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

---

**Anmerkungen**

Die Übereinstimmung mit dem Studienplan ist die Basis für die Nutzbarkeit der Registerdaten. Für eine verfeinerte Analyse könnten die Abweichungen vom Studienplan gewichtet werden.

**Zähler**

Anzahl Abweichungen vom Studienplan

**Nenner**

Anzahl überprüfter Patienten/Personen

**Subkategorien**

---

**Berechnung**

1) Festlegung der zu überprüfenden wesentlichen Elemente des Studienplans und der Verfahren zur Messung von Abweichungen (z.B. bei Einhaltung der Zeitintervalle für Untersuchungen die Angabe einer akzeptablen Zeitdifferenz zwischen geplanter und tatsächlicher Untersuchung)

2) Bestimmung aller Abweichungen vom Studienplan

3) Berechnung von Zähler, Nenner und Rate

Alternative Berechnungsmöglichkeit:

Mittlere Anzahl von Abweichungen vom Studienplan pro Patient/Person

**Interpretation**

Je höher die Rate, desto schlechter ist die Datenqualität. Von der geplanten Datennutzung hängt es ab, welches Ausmaß an Abweichungen akzeptabel ist.

| Einflussfaktoren |
|---|
| Motivation und Qualifikation des Datenerhebungspersonals; Motivation der Patienten/Personen; Form und Inhalt des Studienplans; klare und eindeutige Beschreibung aller Datenelemente; klare und eindeutige Regeln für die Datenerhebung und -erfassung; Qualität der technischen Infrastruktur (automatisches Meldeverfahren zur Wiedereinbestellung, Möglichkeit zur Erfassung von Adresswechseln, Meldung von Zentrumswechseln) |



## II.2.3.6 Repräsentativität der Registerdaten

**Beschreibung**

Repräsentativität der Registerdaten

**Begriffsdefinitionen**

*Repräsentativität:* Die im Register erfasste Population ist eine Stichprobe der zu untersuchenden Gesamtpopulation, die in den wesentlichen Einflussgrößen mit der Gesamtpopulation weitgehend übereinstimmt.

**ID**

TMF-1048

**Typ**

Ergebnisqualität

**Quellen**

Projektpartner

**Bezug**

Die Berechnung erfolgt für einen Datenbestand.

**Alternative Definition**

---

**Anmerkungen**

Die Repräsentativität ist wichtig für die Verallgemeinerbarkeit von Auswertungsergebnissen. Allerdings sind aus der Literatur keine Kriterien bekannt, welche eine objektive Messung der Übertragbarkeit der aus Registerdaten gewonnenen Ergebnisse auf die entsprechende Grundgesamtheit erlauben.

Der Indikator ist nur relevant für Register, die nicht die gesamte Zielpopulation, sondern nur einen repräsentativen Teil der Zielpopulation erfassen sollen.

Bei der Festlegung der wesentlichen Einflussgrößen können anderweitig bekannte Inzidenzen oder Prävalenzen herangezogen werden (z.B. der Anteil drogensüchtiger HIV-Infizierter an der Gesamtzahl der HIV-Infizierten).

**Zähler**

Anzahl der Einflussgrößen mit erwarteter Verteilung

**Nenner**

Anzahl überprüfter Einflussgrößen

**Subkategorien**

---

**Berechnung**

1) Auswahl der für die Repräsentativität wesentlichen Einflussgrößen

2) Bestimmung der Einflussgrößen für die Population des Registers

3) Vergleich der Registerpopulation mit der Gesamtpopulation

4) Berechnung von Zähler, Nenner und Rate

| Interpretation |
| --- |
| Je höher die Rate, desto besser ist die Repräsentativität und damit die Datenqualität. Welche Rate noch akzeptabel ist, hängt von der vorgesehenen Datennutzung ab. |
| **Einflussfaktoren** |
| Art der Zentren (z.B. Krankenhaus, niedergelassener Arzt); räumliche Verteilung der Zentren (z.B. Großstadt, Kleinstadt, Dorf); unterschiedliche Patientenkollektive in verschiedenen Zentren |

### II.2.4 Schwellenwerte

Üblicherweise werden bei ratenbasierten Indikatoren Schwellenwerte festgelegt, mit denen auffällige von unauffälligen Werten unterschieden werden. Dieses Vorgehen wird z.B. von der JCAHO empfohlen [JCAHO 1990]. Bekannte Indikatorensets im Gesundheitswesen wie die Qualitätsindikatoren der Agency for Healthcare Research and Quality (AHRQ)[6] geben durchgehend Schwellenwerte an. Auch zur Anwendung dieser Leitlinie ist daher ein individueller Schwellenwert pro Indikator für die Unterscheidung von (aus Sicht der Datenqualität) „auffälligen“ und „unauffälligen“ Indikatorwerten erforderlich. Dies ist die Voraussetzung, damit aus den einzelnen Indikatorwerten ein Gesamtscore zur Beurteilung der Datenqualität (s. Kap. II.3) gebildet werden kann. Bei der Literatursichtung wurden keine Evidenz-basierten Empfehlungen für Schwellenwerte gefunden. Angesichts der deutlichen Ausweitung der Indikatoren wird in Version 2.0 der Leitlinie auf die Angabe von erfahrungsbasierten Vorschlägen für Schwellenwerte verzichtet. Diese finden sich zum Teil in den Anwendungsempfehlungen für Register, Kohorten und Data Repositories in Kapitel VI. Bei Anwendung dieser Leitlinie ist es daher regelhaft erforderlich, vorhabensspezifische Schwellenwerte zu formulieren.

## II.3 Qualitätsscore

Zur Steuerung der Empfehlungen müssen die Qualitätsindikatoren zu einem Merkmal zusammengefasst werden. In der Medizin bedient man sich hierzu häufig eines Scores. Holle 1995 definiert einen klinischen Score wie folgt:

> *Ein klinischer Score ist eine Messvorschrift für ein Merkmal aus dem Bereich der klinischen Medizin, die aus einer genau operationalisierten Zusammenfassung von mehreren (i.d.R. mehr als zwei) Komponenten besteht und zu einer mindestens ordinalen Bewertung führt. [Holle 1995, S. 4]*

Diese Definition lässt sich bei Weglassen des Adjektivs *klinisch* ohne Schwierigkeiten auf die gegebene Fragestellung übertragen. Bei der Bewertung eines Datenbestandes handelt es sich weiterhin um eine Zustandsbeschreibung und somit um einen evaluativen Score, in Abgrenzung zum diskriminativen und prädiktiven Score als weitere von Holle beschriebene Varianten. Die Anwendung von Scores zur Beschreibung von Datenqualität wurde vereinzelt beschrieben [Tolonen et al. 2006].

Die ersten Schritte zur Bildung des Qualitätsscores sind mit der Identifikation und Beschreibung der Qualitätsindikatoren abgeschlossen. Hiermit sind die Komponenten des Scores festgelegt. Ein nächster Schritt besteht in der Fest-

6 s. Agency for Healthcare Research and Quality (AHRQ) Quality Indicators unter http://www.qualityindicators.ahrq.gov/

legung einer geeigneten Kombination der Komponenten, die z.B. im Rahmen einer multivariaten Regressionsanalyse empirisch abzuleiten wäre. Da zum Zeitpunkt der Leitlinienerstellung weder ein geeigneter Datenbestand noch ein externes Kriterium zur Datenqualität zur Verfügung stand, wurden Gewichte für die einzelnen Qualitätsindikatoren normativ-analytisch festgelegt. Abschließend bleibt noch die Notwendigkeit zur Festlegung eines Verfahrens zur Berechnung des Scorewertes. Auch hier wird dem in Holle 1995 beschriebenen Konzept der gewichteten Summe gefolgt und so die kategorielle, qualitative Eingangsinformation der einzelnen Qualitätsindikatoren (auffällig/unauffällig) auf ein pseudo-quantitatives Niveau angehoben. Hierbei ist zu berücksichtigen, dass nicht in jedem Falle alle Komponenten (= Qualitätsindikatoren) zur Verfügung stehen. Regelhaft ist dies bei der ersten Analyse für die Indikatoren aus der SDV der Fall, da diese noch nicht durchgeführt wurde. Über eine Transformation sollte daher sichergestellt werden, dass der Gesamt-Scorewert unabhängig von der Anzahl der bestimmten Qualitätsindikatoren immer die gleichen Zahlenwerte einnehmen kann. Im letzten Schritt sollte der Gesamt-Scorewert für die Steuerung der Empfehlungen auf eine kategorielle Bewertung der Datenqualität zurückgeführt werden. Die erforderliche Untersuchung von Reliabilität und Validität des Scores bleibt nachfolgenden Studien vorbehalten.

Den beispielhaft angegebenen Gewichten liegt nun die Überlegung zugrunde, dass die Indikatoren der Ebene Richtigkeit die größte Relevanz besitzen, diejenigen der Ebene Organisation eine mittlere und die Indikatoren der Ebene Integrität die geringste. Als zulässige Gewichte wurden ganzzahlige Werte zwischen 1 und 9 festgelegt, um einerseits eine gleichmäßige Verteilung der Werte auf die Ebenen und andererseits eine Differenzierung zwischen Indikatoren der gleichen Ebene zu ermöglichen. Mit der Voraussetzung, dass das geringstmögliche Gewicht den Wert 1 erhält, werden den Indikatoren der Ebene Richtigkeit Werte im Bereich von 7 bis 9 zugeteilt. Die Indikatoren der anderen Ebenen nehmen dann Werte im Bereich von 4 bis 6 (Ebene Organisation) bzw. 1 bis 3 (Ebene Integrität) ein. Entsprechend dem Projektansatz sollten die Indikatoren aus der SDV das höchste Gewicht erhalten, um Situationen auszuschließen, bei denen Qualitätsscore und Übereinstimmung bei der SDV zu stark abweichen. Innerhalb der Ebenen erfolgte dann eine weitere Differenzierung nach Experteneinschätzung. Tabelle 3 zeigt die beispielhaft angegebenen Gewichte im Überblick.

Der Qualitätsscore wird wie folgt ermittelt:

1. Bewertung einer Rate im Vergleich mit dem Schwellenwert des Indikators als auffällig oder unauffällig. Liegt die Rate im unauffälligen Bereich erhält der Indikator den Rohwert 1, ansonsten 0. Für den Vergleich mit dem Schwellenwert werden zwei Alternativen angeboten:
   - Vergleich der berechneten Rate mit dem vorab festgelegten Schwellenwert

- Berechnung eines Intervalls, welches die wahre Rate mit 95%iger Wahrscheinlichkeit umschließt. Das Vorgehen zur Berechnung dieses Intervalls ist in Kapitel II.3.1 beschrieben. Eine Rate wird dann als auffällig gewertet, wenn sowohl untere als auch obere Grenze des 95%-Konfidenzintervalls im Vergleich mit dem vorab festgelegten Schwellenwert auffällig sind.

2. Der Rohwert wird für jeden Indikator mit seinem Gewicht multipliziert. Es ergibt sich der Teil-Scorewert (ST).
3. Bildung der Summe aller Teil-Scorewerte (ST) der bestimmten Qualitätsindikatoren.
4. Bildung der Summe aller Gewichte (SG) der bestimmten Qualitätsindikatoren.
5. Standardisierung durch Transformation auf eine Skala zwischen 0 und 100: Gesamt-Scorewert = 100 x (ST/SG).

Tab. 3 Beispiele für relative Gewichte der Qualitätsindikatoren

| Qualitätsindikator | Gewicht |
|---|---|
| **Ebene Integrität** | |
| Übereinstimmung mit Vorwert (TMF-1001) | 1 |
| Konkordanz (TMF-1002) | 1 |
| Widerspruchsfreiheit (TMF-1003) | 3 |
| ■ Endlosüberlebende (TMF-1035) | 2 |
| ■ Sichere Widersprüche/Fehler (TMF-1004) | 2 |
| ■ Mögliche Widersprüche/Warnungen (TMF-1005) | 1 |
| Werteverteilung (TMF-1006) | 1 |
| ■ Bevorzugung bestimmter Endziffern (TMF-1007) | 1 |
| ■ Werteverteilung der durch Untersucher erfassten Parameter (TMF-1009) | 1 |
| ■ Werteverteilung der durch Geräte erfassten Parameter (TMF-1010) | 2 |
| ■ Werteverteilung von Befunden (TMF-1011) | 1 |
| ■ Werteverteilung von Parametern zwischen Zentren (TMF-1052) | 2 |
| Untersuchungen am Wochenende (TMF-1008) | 1 |
| Fehlende Einträge | |
| ■ Fehlende Module (TMF-1012) | 1 |
| ■ Fehlende Werte bei Datenelementen (TMF-1013) | 2 |
| ■ Fehlende Werte bei mandatorischen Datenelementen (TMF-1014) | 3 |
| ■ Fehlende Werte bei optionalen Datenelementen (TMF-1015) | 1 |
| ■ Datenelemente mit Wert unbekannt o.ä. (TMF-1016) | 1 |
| ■ Datenelemente mit bestehenden Einträgen bei allen Beobachtungseinheiten (TMF-1017) | 2 |

| Qualitätsindikator | Gewicht |
|---|---|
| Ausreißer bei stetigen Datenelementen (TMF-1018) | 2 |
| Werte, die die Messbarkeitsgrenzen von Verfahren unter- oder überschreiten (TMF-1019) | 1 |
| Werte aus Standards (TMF-1020) | 2 |
| Unerlaubte Werte | |
| ■ Unerlaubte Werte bei qualitativen Datenelementen (TMF-1021) | 1 |
| ■ Unerlaubte Werte bei qualitativen Datenelementen zur Kodierung von Missings (TMF-1022) | 1 |
| ■ Unerlaubte Werte zur Kodierung von fehlenden Modulen (TMF-1023) | 1 |
| ■ Unerlaubte Werte bei quantitativen Datenelementen zur Kodierung von Unter- oder Überschreitungen von Messbarkeitsgrenzen (TMF-1024) | 1 |
| Datenelemente mit unspezifischen Werten (TMF-1025) | 1 |
| Beobachtungseinheiten mit unbekanntem Primärtumor (TMF-1026) | 2 |
| Nachweis bekannter Korrelationen (TMF-1027) | 1 |
| Umfang der Metadaten bei Untersuchungen (TMF-1050) | 1 |
| **Ebene Organisation** | |
| Aktualität der gespeicherten Daten (TMF-1028) | 6 |
| Dubletten im Datenbestand (TMF-1029) | 4 |
| Rekrutierungsrate (TMF-1030) | 5 |
| ■ DCO-Rate (Death Certificate Only) (TMF-1051) | 6 |
| Verweigerungsraten | |
| ■ Verweigerungsrate von Untersuchungen (TMF-1031) | 4 |
| ■ Verweigerungsrate von Modulen (TMF-1032) | 4 |
| ■ Verweigerungsrate bei einzelnen Datenelementen (TMF-1033) | 4 |
| Vorzeitig aus dem Register ausscheidende Beobachtungseinheiten (TMF-1034) | 5 |
| Synonyme (TMF-1036) | 4 |
| Homonyme (TMF-1037) | 4 |
| Einzelmeldungen pro Beobachtungseinheit (TMF-1038) | 4 |
| Solitäre Meldungen von Pathologen (TMF-1039) | 5 |
| Zurückgewiesene Meldungen (TMF-1040) | 5 |
| Datenquellen pro Beobachtungseinheit (TMF-1041) | 4 |
| Beobachtungseinheiten mit Follow-up (TMF-1042) | 6 |
| **Ebene Richtigkeit** | |
| Genauigkeit der Registerdaten (TMF-1043) | 8 |
| Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Datenelemente (TMF-1044) | 9 |



| Qualitätsindikator | Gewicht |
|---|---|
| Übereinstimmung der Registerdaten mit den Originaldaten bezogen auf Beobachtungseinheiten (TMF-1045) | 8 |
| Vollständigkeit der Registerdaten (TMF-1046) | 7 |
| Übereinstimmung mit Verfahrensregeln (TMF-1047) | 7 |
| Repräsentativität der Registerdaten (TMF-1048) | 7 |

Bei der Bildung des Gesamt-Scorewertes beeinflusst die Anzahl von Indikatoren einer Ebene die Bedeutung dieser Ebene für die Qualitätsbewertung. So liegt die Summe der Gewichte bei Verwendung aller Indikatoren bei 43 (Integrität), 70 (Organisation) und 46 (Richtigkeit). Auch bei Auswahl einer Teilmenge von Indikatoren muss mit wechselnden Bedeutungen der Ebenen für den Gesamt-Scorewert gerechnet werden. Die Anzahl von Indikatoren sowie deren Verteilung auf die Ebenen ist daher bei der Interpretation eines Gesamt-Scorewertes zu berücksichtigen. Ggf. kann die getrennte Berechnung von Teil-Scorewerten für die einzelnen Ebenen sinnvoll sein.

Bei der Zuordnung der möglichen Zahlenwerte von 0 bis 100 des standardisierten Gesamt-Scorewertes zu einem qualitativen Merkmal wurde eine Einteilung in fünf Kategorien mit proportionaler Zuteilung der Werte gewählt. Tabelle 4 zeigt die sich ergebende Einteilung.

Tab. 4 Einteilung der Kategorien des Qualitätsscores

| Gesamt-Scorewert | Beurteilung |
|---|---|
| 0 bis unter 20 | sehr schlecht |
| 20 bis unter 40 | schlecht |
| 40 bis unter 60 | moderat |
| 60 bis unter 80 | gut |
| 80 bis 100 | sehr gut |

Angesichts der Zunahme von Indikatoren in Version 2.0 der Leitlinie sind die Gewichte und das beschriebene Vorgehen zu ihrer Verdichtung nur beispielhaft zu verstehen. Es ist davon auszugehen, dass nur in ganz seltenen Fällen alle Indikatoren gleichzeitig in einem Vorhaben zur Anwendung kommen. Bei einer Auswahl von Indikatoren sind dann jedoch die Gewichte zu überprüfen und ggf. anzupassen. Anregungen für eine von der Art des Vorhabens abhängige Auswahl und Gewichtung von Indikatoren finden sich in den Anwendungsempfehlungen zu Registern, Kohorten und Data Repositories in Kapitel VI. Für ein Vorhaben sind damit die Angaben in Tabelle 5 festzulegen: die Auswahl der berücksichtigten Indikatoren sowie deren Schwellenwerte und Gewichte.

Tab. 5 Vorhabensbezogene Eigenschaften der Qualitätsindikatoren

| Qualitätsindikator | Auswahl (ja/nein) | Schwellenwert | Gewicht |
|---|---|---|---|
| **Ebene Integrität** | | | |
| Übereinstimmung mit Vorwert (TMF-1001) | | | |
| Konkordanz (TMF-1002) | | | |
| Widerspruchsfreiheit (TMF-1003) | | | |
| … | | | |

## II.3.1 Bildung eines Intervalls um die wahre Rate (95%-Konfidenzintervall)

Das Intervall um die wahre Rate wird wie von Hartung et al. [1995] angegeben unter Verwendung der F-Verteilung mit Hilfe der Pearson-Clopper-Werte ermittelt. Die Vorgehensweise ist für kleine Fallzahlen geeignet. Die Irrtumswahrscheinlichkeit wird mit 0,05 angesetzt (= α). Gleichung (1) zeigt die Formel zur Berechnung der unteren Grenze $\mathbf{p_1}$ für das 95%-Konfidenzintervall der wahren Rate, Gleichung (2) die Formel zur Berechnung der oberen Grenze $\mathbf{p_2}$.

$$p_1 = \frac{mF_{2m,2(n-m+1);\alpha/2}}{n-m+1+mF_{2m,2(n-m+1);\alpha/2}} \tag{1}$$

$$p_2 = \frac{(m+1)F_{2(m+1),2(n-m);1-\alpha/2}}{n-m+(m+1)F_{2(m+1),2(n-m);1-\alpha/2}} \tag{2}$$

Als **n** ist jeweils die als Nenner ermittelte Zahl, als **m** die als Zähler ermittelte Zahl einzusetzen. Falls keine Ereignisse eingetreten sind, also **m = 0**, ist $\mathbf{p_1 = 0}$. Falls Zähler und Nenner gleich sind, also **m = n**, ist $\mathbf{p_2 = 1}$. $\mathbf{p_1}$ und $\mathbf{p_2}$ sind beides Raten, keine absoluten Häufigkeiten. In der Berechnung von $\mathbf{p_1}$ gilt **α/2 = 0,025**, in der Berechnung von $\mathbf{p_2}$ gilt **1 – α/2 = 0,975**.

Anhang 1 enthält eine Tabelle mit Kombinationen von Nenner und Zähler. Folgende Werte wurden hierbei für **n** berücksichtigt: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000, 10.000. **m** wurde für alle **n** über die Raten zwischen 0 und 0,5 in Schritten von 0,01 ermittelt. Nicht ganzzahlige Werte wurden auf den nächsten ganzzahligen Wert aufgerundet. Hieraus ergaben sich 1928 unterschiedliche Kombinationen. Das F-Quantil wurde mit der Inversen Verteilungsfunktionen von SPSS® 12.0.1 ermittelt. Die Werte für $\mathbf{p_1}$ wurden abschließend auf 2 Nachkommastellen abgerundet, die Werte für $\mathbf{p_2}$ auf 2 Nachkommastellen aufgerundet.



## II.4 Metadaten von empirischen Forschungsvorhaben

Zusätzlich zur Datenqualität sollen bei den Empfehlungen weitere Charakteristika eines Vorhabens berücksichtigt werden. Einerseits können diese ebenfalls zur Steuerung von SDV und Feedback beitragen. So ist selbstverständlich eine Fallzahlplanung für eine SDV unsinnig, falls Originaldaten gar nicht zur Verfügung stehen. Anderseits handelt es sich um Informationen, die sinnvollerweise Bestandteil eines Berichtswesen sind, wie z.B. der Name des Vorhabens oder Bezeichnungen der Erhebungszentren. Diese vorhabensspezifischen Charakteristika werden als Metadaten, also Daten über Daten [Berners-Lee 1997] bezeichnet.

Für Version 1.0 der Leitlinie wurden Vorschläge vom IMIBE und den Projektpartnern konsekutiv in den Projektsitzungen eingebracht und über das Protokoll zum Review zurückgespielt. Für eine Operationalisierung der Vorschläge wurde ein Rückgriff auf internationale Standards oder Quasi-Standards angestrebt. In Frage kommende Konzepte wurden daher auf Abdeckung der konsentierten Merkmale geprüft. Die erforderlichen Dokumente wurden über das World Wide Web bezogen. In diese Prüfung wurden folgende Konzepte eingeschlossen:

- Clinical Data Interchange Standards Consortium (CDISC) (http://www.cdisc.org/)
  - Study Data Tabulation Model (SDTM), Version 1.0, 25.6.2004
  - Operational Data Model (ODM), Version 1.2, 19. Dezember 2003
  - Protocol Elements for Regulated Clinical Trials, V2.0, 1. August 2005
- ClinicalTrials.gov (http://www.clinicaltrials.gov/)
  - Data Element Definitions, Dezember 2005
- European Clinical Trials Database (http://eudract.emea.eu.int/)
  - Clinical Trial Application Form, EUDRACT User Manual EUD 134, Version 1.3, April 2004
- Health Level 7 (HL7) (www.hl7.org)
  - HL7 Structured Clinical Trial Protocol, Release 1.0
- IMIBE
  - Standard Operating Procedure Prüfplan
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (http://www.ich.org/)
  - E3 – Structure and Content of Clinical Study Reports
  - E6 – Good Clinical Practice
  - Common Technical Document (CTD)

Alleinig die Protocol Elements for Regulated Clinical Trials von CDISC (CDISC – PR) zeigten eine ausreichende Abdeckung der konsentierten Metadaten. Allerdings fehlt noch eine formale Repräsentation, wie sie z.B. beim HL7 Structured Clinical Trial Protocol, Release 1.0 in Form eines XML-Schemas vorgelegt wird. Es wurde daher entschieden, wo möglich, die Metadaten auf CDISC – PR zurückzuführen, weitere Definitionen wie Wertebereiche dort, wo nötig, aus den Projektarbeiten abzuleiten.

Tabelle 6 zeigt die Metadaten mit ihrer Bezeichnung, der Art ihrer Verwendung innerhalb der Leitlinie, ggf. geeignete Merkmale bei CDISC – PR sowie den Wertebereich.

Tab. 6 Metadaten von empirischen Forschungsvorhaben

| Bezeichnung | Art der Verwendung | CDISC – PR | Wertebereich |
|---|---|---|---|
| **Allgemeine Angaben** | | | |
| Bezeichnung | Bericht | | Neu: Text |
| Laufzeit der Kohorte | Empfehlung Follow-up Analyse | 119 Duration of subject participation | CDISC/SDTM: ISO 8601 |
| Follow-Up-Frequenz | Empfehlung Follow-up Analyse | 122 Subject follow up | CDISC/SDTM: ISO 8601 |
| Angestrebtes Niveau der Datenqualität | Bericht | | Neu: entsprechend Score TMF |
| **Datenmanagement** | | | |
| Form der Intervention (SDV, Feedback) | Entscheidung über Art der Empfehlungen | 230 Monitoring of trial<br>(231 Site audits)<br>(235 Site monitoring statement)<br>236 Data quality/ integrity statement | Monitoring der Studie: [SDV ja \| SDV nein]<br>Feedback: [Feedback ja \| Feedback nein]<br>Art des Feedbacks: [nur die eigenen Ergebnisse \| anonymisierter Vergleich \| offener Vergleich] |
| **Ressourcen** | | | |
| Verfügbarkeit der Daten für eine SDV | Entscheidung über Art der Empfehlungen | 255 Source documents | Originaldokumentation: [Originaldokumente verfügbar \| Originaldokumente nicht verfügbar] |
| Kapazität für SDV pro Analysezeitraum | Festlegung Frequenz SDV | | Neu: Anzahl Akten pro Monat (Default – Summe über alle Zentren der Zahl der Patienten im Analysezeitraum) |
| Kapazität für SDV pro Besuch | Festlegung Frequenz SDV | | Neu: Anzahl Akten pro Besuch (Default 30) |
| **Aufwand für SDV** | | | |
| Anzahl der Merkmale je Beobachtungseinheit | Fallzahlplanung Tiefe | | Neu: Integer |
| Je Zentrum | | | |
| ■ Bezeichnung | Bericht | | Neu: Text |
| ■ Zahl der Patienten im Analysezeitraum | Fallzahlplanung Umfang | | Neu: Integer |



# II.5 Planung der Source Data Verification

## II.5.1 Einführung

Die Literaturrecherche zu Version 1.0 der Leitlinie lieferte keine durch empirische Untersuchungen begründeten Empfehlungen für die Durchführung eines Monitorings mit einer auf Stichproben basierenden Source Data Verification. Vereinzelt werden Aussagen zum Umfang der Stichprobe für eine SDV gemacht:

- 5% der wichtigen und 1% der anderen CRF [Pogash et al. 2001]
- Mindestens 10% der Daten pro Zentrum [Califf et al. 1997]
- 10% der Patienten, bei wichtigen Studien bis zu 40% der Patienten [Weiss 1998]
- Ungefähr 10% der in einem Jahr angefallenen Operationen, 1/3 der beteiligten Kliniken [Håkansson et al. 2001]
- Zufällige Auswahl von 13% der Patienten [O'Sullivan et al. 1996]
- 15% bis 20% der Patienten [Weiss et al. 1993]

Daneben gibt es Ansätze, den Umfang der SDV von der Datenqualität abhängig zu machen, d.h. häufigeres Monitoring bei niedriger Qualität und selteneres Monitoring bei hoher Qualität:

- Werden mehr als 10 Abweichungen pro 10.000 Felder gefunden, werden weitere 5% der CRF geprüft [Pogash et al. 2001].
- Bei den ersten 10% der CRF werden alle Datenelemente überprüft. Die aus dieser Stichprobe geschätzte Fehlerrate bestimmt den Umfang der partiellen Kontrolle der restlichen Datenelemente [Mullooly 1990].
- Ose et al. 2004 beschreiben verschiedene Ansätze zu einem adaptiven Monitoring bei klinischen Studien.

In dieser Leitlinie wird ein anderer, formaler Weg zur Bestimmung der Stichprobengröße eingeschlagen, wie er in ähnlicher Weise bei der Fehlbelegungsprüfung im Krankenhausbereich genutzt wird [VdAK et al. o.J.]. Die Stichprobengröße wird an die Datenqualität angepasst. Bei Zentren mit guter Datenqualität fällt die Stichprobe für die SDV kleiner aus als bei Zentren mit schlechter Datenqualität. Die Datenqualität wird anhand des Qualitätsscores und – sofern bereits eine SDV durchgeführt wurde – des Ergebnisses der vorherigen SDV bestimmt. Die Stichprobengröße wird mittels bekannter statistischer Verfahren zur Fallzahlplanung [beschrieben z.B. bei Elashoff 1999] berechnet. Zur vollständigen Planung der SDV gehören die Bestimmung des Umfangs, der Tiefe und der Frequenz der SDV für jedes Zentrum.

## II.5.2 Bestimmung des Umfangs der Source Data Verification

Der Umfang ist definiert als die Anzahl Beobachtungseinheiten (z.B. Personen, Patienten, Behandlungsfälle, Meldungen, Erhebungsbögen) pro Zentrum, bei

der eine SDV durchgeführt werden soll. Diese werden in jedem Zentrum zufällig ausgewählt. Die Fallzahlplanung basiert hier auf der Schätzung des Anteils an Beobachtungseinheiten mit mindestens einem fehlerhaften Eintrag in den Daten. Es kann eine Binomialverteilung angenommen werden, die durch die Normalverteilung approximiert wird. Um stärkere Fallzahlunterschiede zwischen Zentren mit guter und schlechter Datenqualität zu erreichen, wird die geforderte absolute Breite des Konfidenzintervalls in Abhängigkeit vom Wert des Qualitätsscores umso kleiner gewählt, je schlechter die Datenqualität ist.

Für die Fallzahlschätzung müssen zuerst folgende allgemeine Festlegungen getroffen werden:

1. Festlegung der Beobachtungseinheit, die für alle Zentren gleich gewählt werden muss.
2. Festlegung der absoluten halben Breite **δ** des 95%-Konfidenzintervalls für jede Kategorie des Qualitätsscores. Das Konfidenzintervall hat die Form (Schätzwert −**δ**, Schätzwert +**δ**) und liegt symmetrisch um den Schätzwert mit einem Abstand **δ** > 0. Die absolute Breite wird spezifiziert durch die Angabe des Abstandes **δ** des Schätzwertes von den Intervallgrenzen. Tabelle 7 zeigt die empfohlenen Werte. Eine vorhabensspezifische Anpassung an ein einzelnes Register oder eine Kohortenstudie ist möglich. Es ist darauf zu achten, dass **δ** umso kleiner gewählt wird, je schlechter die Datenqualität ist.

Tab. 7 Absolute halbe Breite des Konfidenzintervalls für die Kategorien des Qualitätsscores

| Gesamt-Scorewert | Beurteilung | Empfehlung für δ |
|---|---|---|
| 0 bis unter 20 | sehr schlecht | 0,01 |
| 20 bis unter 40 | schlecht | 0,02 |
| 40 bis unter 60 | moderat | 0,03 |
| 60 bis unter 80 | gut | 0,04 |
| 80 bis 100 | sehr gut | 0,05 |

Danach wird für jedes Zentrum separat wie folgt vorgegangen:

1. Festlegung des Anteils **p** an Beobachtungseinheiten mit mindestens einem fehlerhaften Eintrag in den Daten. Beim 1. Zyklus der Leitlinienanwendung wird der vorab festgelegte Schwellenwert angesetzt (empfohlen ist **p = 0,05**), ab dem 2. Zyklus wird für **p** der Anteil an Beobachtungseinheiten mit mindestens einem fehlerhaften Eintrag in den Daten aus der vorherigen SDV (Qualitätsindikator) verwendet. Hierbei sind zwei Sonderfälle zu berücksichtigen:
   - Falls **p > 0,5**, dann wird gesetzt **p = 0,5**. Mit dieser oberen Grenze wird eine auf Grund der gewählten Binomialverteilung eintretende Verringerung der Fallzahl bei höheren Fehlerraten ausgeschlossen.



- Falls **$p < 0{,}0062$**, dann wird gesetzt **$p = 0{,}0062$**. Mit dieser unteren Grenze wird in jedem Fall eine minimale SDV sichergestellt. Die Häufigkeit von 6.200 Fehlern auf 1 Million Beobachtungseinheiten entspricht dem Industriestandard in der Six Sigma-Methodologie [Lucas 2002].

2. Bestimmung von **$\delta$** in Abhängigkeit vom ermittelten Wert des Qualitätsscores.
3. Berechnung der nicht-adjustierten Fallzahl **$n_0$** nach der Formel in Gleichung (3). Dabei ist $z_{1-\alpha/2}$ das Quantil der Standard-Normalverteilung. Für den hier vorgegebenen Fehler 1. Art **$\alpha = 0.05$** ist $z_{1-\alpha/2} = 1{,}96$.

$$n_0 = \frac{p(1-p)}{\delta^2} \cdot z^2_{1-\alpha/2} \qquad (3)$$

4. Da die im vorherigen Schritt berechnete Fallzahl **$n_0$** größer sein kann als die Gesamtzahl aller Beobachtungseinheiten im Zentrum, ist eine Adjustierung der Fallzahl für die Zentrumsgröße erforderlich. Sie geschieht nach der Formel in Gleichung (4). Hier ist **N** die Gesamtzahl aller Beobachtungseinheiten im Zentrum (angegeben in den Metadaten) und **n** die adjustierte Fallzahl. Ein nicht-ganzzahliger Wert wird hier immer auf den nächsthöheren ganzzahligen Wert aufgerundet. **n** gibt an, wie viele Beobachtungseinheiten bei der nächsten SDV zu überprüfen sind.

$$n = \frac{n_0 \cdot N}{n_0 + N} \qquad (4)$$

**Beispiel**

**$p = 0{,}05$** und **$\delta = 0{,}04$** ergibt eingesetzt in die rechte Seite der Gleichung (5) eine nicht-adjustierte Anzahl von zu prüfenden Beobachtungseinheiten **$n_0 = 114{,}05$**. Sei **$N = 100$**, dann ergibt sich aus Gleichung (6) eine adjustierte Anzahl von zu prüfenden Beobachtungseinheiten **$n = 53{,}28$**, die auf **54** aufgerundet wird. Das heißt: in diesem Fall werden 54 von 100 Beobachtungseinheiten in der SDV überprüft.

Anhang 1 enthält eine Tabelle, in der für viele Konstellationen die Fallzahl direkt abgelesen werden kann. Für den Anteil **p** wurden dort in Schritten von 0,01 alle Werte von 0,01 bis 0,5 sowie als untere Grenze 0,0062 berücksichtigt, für **$\delta$** die Werte aus Tabelle 7. Für **N** wurden folgende Werte angesetzt: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000, 10.000. Die sich ergebenden Werte für **n** wurden aufgerundet.

### II.5.3 Bestimmung der Tiefe der Source Data Verification

Unter der Tiefe der SDV wird hier die Angabe verstanden, wie viele (und welche) Datenelemente pro Beobachtungseinheit überprüft werden sollen. Hier sind zwei Alternativen vorgesehen:

- Auswahl aller Datenelemente
- Formale Fallzahlplanung (Annahme einer Binomialverteilung mit Adjustierung der Fallzahl)

Letzteres wird für jedes Zentrum wie folgt durchgeführt:

1. Festlegung des Anteils **p** an Datenelementen mit einem fehlerhaften Eintrag in den Daten. Beim 1. Zyklus der Leitlinienanwendung wird der vorab festgelegte Schwellenwert angesetzt (empfohlen ist **p = 0,05**), ab dem 2. Zyklus wird für **p** der Anteil an Datenelementen mit einem fehlerhaften Eintrag in den Daten aus der vorherigen SDV (Qualitätsindikator) verwendet. Hierbei sind zwei Sonderfälle zu berücksichtigen:
   - Falls **p > 0,5**, dann wird gesetzt **p = 0,5**. Mit dieser oberen Grenze wird eine auf Grund der gewählten Binomialverteilung eintretende Verringerung der Fallzahl bei höheren Fehlerraten ausgeschlossen.
   - Falls **p < 0,0062**, dann wird gesetzt **p = 0,0062**. Mit dieser unteren Grenze wird in jedem Fall eine minimale SDV sichergestellt. Die Häufigkeit von 6200 Fehlern auf 1 Million Beobachtungseinheiten entspricht dem Industriestandard in der Six Sigma-Methodologie [Lucas 2002].
2. Bestimmung von **δ** in Abhängigkeit vom ermittelten Wert des Qualitätsscores (s. Tab. 7).
3. Berechnung der nicht-adjustierten Fallzahl $\mathbf{n_0}$ nach der Formel in der Gleichung (5). Dabei ist $z_{1-\alpha/2}$ das Quantil der Standard-Normalverteilung. Für den hier vorgegebenen Fehler 1. Art **α = 0.05** ist $z_{1-\alpha/2} = 1{,}96$.

$$n_0 = \frac{p(1-p)}{\delta^2} \cdot z^2_{1-\alpha/2} \qquad (5)$$

4. Da die im vorherigen Schritt berechnete Fallzahl $\mathbf{n_0}$ größer sein kann als die Gesamtzahl aller Datenelemente, ist eine Adjustierung der Fallzahl erforderlich. Sie geschieht nach der Formel in Gleichung (6). Hier ist **N** die Gesamtzahl aller Datenelemente einer Beobachtungseinheit (angegeben in den Metadaten) und **n** die adjustierte Fallzahl. Ein nicht-ganzzahliger Wert wird hier immer auf den nächsthöheren ganzzahligen Wert aufgerundet. **n** gibt an, wie viele Datenelemente jeder Beobachtungseinheit bei der nächsten SDV zu überprüfen sind.

$$n = \frac{n_0 \cdot N}{n_0 + N} \qquad (6)$$



**Beispiel**

**p = 0,05** und **δ = 0,04** ergibt eingesetzt in die rechte Seite der Gleichung (5) eine nicht-adjustierte Anzahl von zu prüfenden Datenelementen (pro Beobachtungseinheit) $\mathbf{n_0}$ **= 114,05**. Sei **N = 100**, dann ergibt sich aus Gleichung (6) eine adjustierte Anzahl von zu prüfenden Datenelementen **n = 53,28**, die auf **54** aufgerundet wird. Das heißt: in diesem Fall werden pro Beobachtungseinheit 54 von 100 Datenelementen in der SDV überprüft.

Die Auswahl der Datenelemente kann zufällig oder nach Wichtigkeit erfolgen. Zum Ablesen der Mindestzahl an Merkmalen pro Beobachtungseinheit kann ebenfalls die Tabelle in Anhang 1 herangezogen werden (s. Kap. II.5.2).

### II.5.4 Bestimmung der Frequenz der Source Data Verification

Die aufgrund der Fallzahlplanung erforderliche SDV sollte unter Berücksichtigung der vorhandenen Ressourcen (Kapazität für SDV pro Analysezeitraum und Besuch sind Teil der Metadaten, s. Kap. II.4) gleichmäßig auf den zur Verfügung stehenden Zeitraum bis zur nächsten Anwendung der Leitlinie aufgeteilt werden.

**Beispiel**

6 Monate Zeit für SDV mit Fallzahl von 139 Patienten

SDV bei 30 Patienten pro Besuch möglich

→ 5 Besuche gleichmäßig auf 6 Monate aufteilen.

## II.6 Planung des Feedbacks

Ohne Information über die eigene Qualität lassen sich keine Defizite erkennen und existieren keine Anreize zur Qualitätsverbesserung [Selbmann 2000]. Daher ist das Feedback neben der SDV als weitere Interventionsmöglichkeit vorgesehen. Jedes an der Datensammlung beteiligte Zentrum erhält einen individuellen Bericht über seine Datenqualität. Der Bericht sollte folgenden Inhalt haben:

- Bezeichnung des Registers/der Kohorte
- Bezeichnung des Zentrums
- angestrebtes Niveau der Datenqualität
- Wert des Qualitätsscores (Scorewert und Kategorie)
- Werte der erhobenen Qualitätsindikatoren
- Schwellenwerte und Gewichtungen der erhobenen Qualitätsindikatoren
- Indikatordefinitionen für Qualitätsindikatoren mit auffälligem Wert

Je Zentrum können alternativ

- nur die eigenen Ergebnisse oder
- die eigenen Ergebnisse zusammen mit anonymisierten Vergleichsdaten (Qualitätsindikatoren, Qualitätsscore) der anderen Zentren oder
- die eigenen Ergebnisse zusammen mit offenen Vergleichsdaten (Qualitätsindikatoren, Qualitätsscore) der anderen Zentren

bekannt gemacht werden. Für jedes Register/jede Kohortenstudie muss individuell entschieden werden, welche Variante bevorzugt wird. Im Kompetenznetz Parkinson findet derzeit eine vergleichende Untersuchung dieser drei Feedback-Strategien statt. Die ersten vorläufigen Ergebnisse erlauben noch keine Angabe einer Präferenz für eine der Feedback-Strategien.

Die Empfehlungen zur Frequenz des Feedbacks variieren in der Literatur sehr stark (von monatlich bis jährlich). Als Minimum sollte ein Feedback immer dann erfolgen, wenn die Werte der Qualitätsindikatoren und des Qualitätsscores im Rahmen der Anwendung dieser Leitlinie neu ermittelt worden sind.

## II.7 Zeitpunkt der nächsten Analyse

Der Zeitpunkt für die nächste Analyse ist mit einem Default von sechs Monaten vorbelegt. Eine Anpassung der Zeitspanne an die Gegebenheiten eines Vorhabens ist vorgesehen. Dabei sollte die Zeitspanne so groß gewählt werden, dass die durchgeführten Maßnahmen (SDV, Feedback) Wirkung entfalten können und von den Zentren genug neue Daten für eine SDV erfasst werden können. Zu berücksichtigen sind speziell die Laufzeit des gesamten Vorhabens, die Frequenz des Follow-ups, die Rekrutierungsrate an neuen Patienten/Personen pro Zentrum und das erwartete Volumen neuer Daten, die von den Zentren eingebracht werden.

# III Ergebnisse

Das adaptive Monitoring verfolgt aufgrund der Stichproben-basierten SDV nicht das Ziel, eine vollständige Übereinstimmung der erhobenen Daten mit der Primärdokumentation sicherzustellen, wie dies beim herkömmlichen Monitoring klinischer Studien der Fall ist. Vielmehr wird im Sinne einer Qualitätslenkung die Beeinflussung der meldenden Einrichtungen angestrebt, um eine primär hohe Datenqualität zu erreichen und zu halten bzw. eine am Anfang des Vorhabens mäßige oder schlechte Datenqualität nachhaltig positiv zu beeinflussen. Die für ein Register oder eine Kohortenstudie verfügbaren Ressourcen werden in den Bereichen konzentriert eingesetzt, wo sie am effektivsten wirken können. Hierzu werden dem Verantwortlichen für das Datenmanagement folgende Ergebnisse zur Verfügung gestellt:

Für jedes Zentrum die Mindestzahl an Beobachtungseinheiten, für die eine SDV durchzuführen ist (Umfang der SDV).

- Für jedes Zentrum die Mindestzahl an Merkmalen, die für jede Beobachtungseinheit bei der SDV zu bearbeiten ist (Tiefe der SDV).
- Für jedes Zentrum die Anzahl an Besuchen für die SDV im Zyklus, entsprechend der angegebenen Kapazität (Frequenz der SDV).
- Für jedes Zentrum die Raten der einzelnen (maximal 24) Qualitätsindikatoren mit Bewertung in auffällig/unauffällig im Vergleich zum Schwellenwert.
- Für jedes Zentrum der Qualitätsscore.

# IV Wissenschaftliche Belege

Die Ergebnisse der Literatursichtung sind detailliert in den Anhängen 2 und 3 dargestellt. Hier sollen nur einige wichtige Ergebnisse vorgestellt werden. Grundsätzlich hat sich gezeigt, dass eine empirische Absicherung für Strategien zum Datenmanagement und Monitoring im Bereich von Kohortenstudien und Registern bislang weitgehend fehlt. Ähnliches scheint für klinische Studien zu gelten. Die in der Industrie zum Qualitätsmanagement angewandten Modelle (z.B. Total Quality Management) erscheinen für das gegebene Problem zu komplex.

Die für ein Register oder eine Kohortenstudie notwendige und erreichbare Datenqualität hängt vor allem von der jeweiligen Fragestellung bzw. der geplanten Nutzung der Daten ab sowie von den Prozessen, die die Daten erzeugen. Die Motivation der Beteiligten hat Einfluss auf die Datenqualität und steigt, wenn den Datenlieferanten selbst Vorteile aus guter bzw. verbesserter Datenqualität entstehen. Das Monitoring ist hilfreich bei der Weckung und Schärfung des Qualitätsbewusstseins. Allerdings sind nur wenige quantitative Daten zum Monitoring klinischer Studien in der Literatur verfügbar. Die Source Data Verification (SDV) wird oft zur Untersuchung von Vollständigkeit und/oder Genauigkeit bei klinischen Studien, Registern und Elektronischen Patientenakten eingesetzt. Es finden sich einzelne Empfehlungen (z.B. zu Frequenz, Anzahl Variablen, Interpretation der Nicht-Übereinstimmungen), die allerdings nicht durch empirische Untersuchungen mit Vergleich unterschiedlicher Ansätze begründet sind. Eine Untersuchung der Datenqualität in Registern geschieht meist durch Vergleich mit anderen Datenquellen. Der Goldstandard für den Vergleich ist unklar, die Patientenakte ist dafür nur bedingt geeignet. Eine Reihe von Indikatoren für Datenqualität werden in der Literatur beschrieben. Manche werden mehrfach von unterschiedlichen Autoren genannt. Eine zeitnahe Rückmeldung der erreichten Datenqualität an die Datenlieferanten macht Defizite sichtbar und gibt Impulse zur Qualitätsverbesserung. Ein (anonymisiertes) Benchmarking kann diesen Effekt verstärken.

# V Nutzen

Bei Anwendung der Leitlinie entsteht ein Nutzen in zweifacher Hinsicht. Zum einen bietet die Leitlinie ein Konzept zur kontinuierlichen Verbesserung von Datenqualität. Hierdurch kann die Aussagefähigkeit und Nutzbarkeit der Daten von Kohortenstudien und Registern wesentlich gestärkt werden. Zum anderen wird ein Originaldatenabgleich (Source Data Verification, SDV) im Rahmen von Stichproben empfohlen, wodurch eine deutliche Reduktion von Aufwand beim Monitoring im Vergleich zur vollständigen SDV erreicht wird. Es entsteht also Zusatznutzen bei gleichzeitiger Aufwandsreduktion.

Die Anwendung der Leitlinie erfordert eine konsequente und systematische Auseinandersetzung mit der Qualität der Daten im zentralen Datenbestand des jeweiligen Vorhabens. Im Ausnahmefall, der Nutzung aller Qualitätsindikatoren, sind 102 Parameter (d.h. Zähler und Nenner der Qualitätsindikatoren), zu deren Berechnung wiederum zum Teil eine gesonderte Analyse der Daten erforderlich ist, zur Berechnung von 51 Qualitätsindikatoren zu erheben. Die folgenden Verfahrensschritte – im Wesentlichen zur Planung der SDV – sind dann mit wenig Aufwand durchzuführen. Die Leitlinie räumt aber auch Spielraum in diesem Verfahrensschritt ein, indem mit einem vorhabensspezifischen Teilbestand an Qualitätsindikatoren gearbeitet werden kann. Qualitätsindikatoren, deren Ermittlung im einzelnen Fall einen unverhältnismäßig hohen Aufwand bedeuten würde, können weggelassen werden. Die Verantwortung hierfür liegt allerdings beim Vorhaben selbst.

Durch die Anpassung der Maßnahmen an die Datenqualität wird der verbleibende Aufwand, im Wesentlichen für die SDV, dort konzentriert, wo es besonders nötig ist. Bei erfolgreicher Anwendung über mehrere Zyklen ist automatisch mit einer weiteren Aufwandsreduktion im Monitoring bei sich verbessernder Datenqualität zu rechnen.

Für Vorhaben, die bislang weder eine Analyse der Datenqualität noch eine SDV oder ein Berichtswesen durchgeführt haben, kann dennoch in der Summe ein nicht unbeträchtlicher Aufwand entstehen. Ein Verzicht auf alle genannten Maßnahmen führt allerdings aus heutiger Sicht zu einem nicht verwertbaren

Datenbestand, so dass diese Situation für einen Vergleich des Aufwandes irrelevant ist.

Zusammenfassend lässt sich der Nutzen wie folgt charakterisieren:

- Die Datenqualität der Erhebungszentren und des zentralen Datenbestandes wird bestimmbar.
- Eine vollständige und sehr aufwendige SDV wird vermieden.
- Die Ressourcen werden dort konzentriert, wo sie auf Grund schlechterer Datenqualität besonders nötig sind.
- Es entsteht ein Prozess der kontinuierlichen Verbesserung von Datenqualität.

# VI Empfehlungen

## VI.1 Allgemeine Empfehlungen

Eine hohe Datenqualität ist für die Aussagekraft und Nutzbarkeit von Kohortenstudien und Registern von entscheidender Bedeutung. Die Leitlinie empfiehlt hierzu die Durchführung eines Originaldatenabgleiches (Source Data Verification, SDV) einerseits und die Implementierung eines Berichtswesens zur Datenqualität (Feedback) andererseits. Die Leitlinie sollte regelmäßig angewendet werden. Als Abstand zwischen zwei Zyklen werden sechs Monate empfohlen.

Mit dem Ziel der kontinuierlichen Verbesserung der Datenqualität wird eine SDV in Stichproben von Beobachtungseinheiten als ausreichend angesehen. Hierdurch können relevante Ressourcen im Monitoring eingespart werden. Umfang (Anzahl der Beobachtungseinheiten, für die eine SDV durchgeführt wird), Tiefe (Anzahl der Merkmale, die je Beobachtungseinheit abgeglichen werden) und Frequenz (Häufigkeit von Besuchen vor Ort) sind für jedes Zentrum an die Datenqualität anzupassen. Zur Bestimmung der Datenqualität definiert diese Leitlinie 51 Indikatoren und führt diese zu einem Score mit fünf Kategorien (sehr gut – gut – moderat – schlecht – sehr schlecht) zusammen. Über den Score sowie die Vorergebnisse bei der SDV werden Umfang und Tiefe der SDV für jedes Zentrum an die Datenqualität angepasst. Gemäß den Empfehlungen ist dann eine SDV durchzuführen, die einerseits qualitätslenkend auf die lokalen Erhebungszentren einwirkt und andererseits dem Aufdecken von Schwachstellen bei der Datenerfassung und -weitergabe aber auch im Studiendesign dient. Eine Korrektur von Daten, die bei der SDV als fehlerhaft erkannt wurden, ist nicht Ziel der Leitlinie. Diese wird nur dann empfohlen, wenn hierzu der gesamte Datenbestand – ggf. über die Verwendung zusätzlicher, für das gegebene Problem definierter Plausibilitätsprüfungen – herangezogen wird.

Das Berichtswesen dient ebenfalls der kontinuierlichen Verbesserung der Datenqualität, in dem es motivierend auf lokale Erhebungszentren einwirkt,

die im Vergleich mit anderen eine schwächere Datenqualität aufweisen. Gleichermaßen wie die SDV sind die erreichten Raten der Qualitätsindikatoren und der Qualitätsscore für jedes Zentrum, aber auch für ein Gesamtvorhaben, Grundlage einer Schwachstellenanalyse. Ein Berichtswesen sollte daher ebenfalls systematisch und regelmäßig durchgeführt werden.

## VI.2 Anwendungsempfehlungen für Register

Ulrike Bauer

### VI.2.1 Einleitung

Der Begriff des Registers wird im Forschungsbereich auf vielfältige Weise verwendet. Je nach verfolgtem Ziel gibt es unterschiedliche Arten von Registern:

- bevölkerungsbezogene, epidemiologische
- und klinische, krankheitsbezogene Register.

Insgesamt lassen sich Register beschreiben als eine „möglichst aktive, standardisierte Dokumentation von Beobachtungseinheiten zu vorab festgelegten, aber im Zeitverlauf erweiterbaren Fragestellungen, für die ein präziser Bezug zur Zielpopulation transparent darstellbar ist“ (Müller et al. 2010: S. 825). Bezüglich der Planung, Implementierung und Datenqualität liefern die „Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis“ der Deutschen Arbeitsgemeinschaft für Epidemiologie allgemein gehaltene Hinweise. Diese können als wertvoller Orientierungsrahmen genutzt werden. Die detaillierte Ausgestaltung von Maßnahmen zur Qualitätssicherung hängt sodann vom jeweiligen Register ab.

Bei der Planung eines Registers ist es zunächst unabdingbar, zu identifizieren, um welche Art von Register es sich handeln soll (s.o.).

Es soll beispielhaft am Nationalen Register für angeborene Herzfehler ein flächendeckendes, offenes klinisches Patientenregister dargestellt werden; dabei soll gezeigt werden, welche Maßnahmen zur optimalen Gestaltung der Datenqualität hierbei eingesetzt werden.

Im Nationalen Register für angeborene Herzfehler werden personenbezogen Patienten aller Altersgruppen mit jeglicher Art einer angeborenen Herzerkrankung ohne vorgegebene Endpunkte erfasst. Diese Fehlbildungen sind die häufigste angeborene Erkrankung beim Menschen. Es handelt sich um eine Vielzahl an klinischen Erscheinungsformen mit unterschiedlichsten Ausprägungen, die 70 ICD-10-Codes umfasst, welche pathomorphologisch mit 206 IPC-Codes abgebildet werden. Einzelne Entitäten treten deshalb zum Teil nur selten auf, sodass sich ausreichende Fallzahlen für aussagekräftige Studien oft nur schwer erreichen lassen.




Das seit 2000 bestehende und vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen des Kompetenznetzes Angeborene Herzfehler seit 2003 geförderte Nationale Register für angeborene Herzfehler e.V. erfasst deutschlandweit Daten und Biomaterial zu Personen mit angeborenen Herzfehlern. Bis 2013 wurden rund 50.000 Patienten erfasst.

Durch die zentrale Sammlung und Bündelung von Daten möchte das Nationale Register der Problematik zu geringer Fallzahlen in einzelnen Kliniken begegnen und so eine Grundlage für effiziente Forschung in verschiedenen Bereichen der klinischen, epidemiologischen und genetischen Forschung der angeborenen Herzfehler bereit stellen. Drauf aufbauend können auch spezifische Kohorten für einzelne Studienvorhaben gebildet werden.

## VI.2.2 Schwerpunkte der Qualitätssicherung in Registern

Damit sie als Grundlage für verschiedene Studien mit validen Ergebnissen herangezogen werden können, ist die Qualität der erfassten Daten in Registern von zentraler Bedeutung. Im Gegensatz zu Kohortenstudien werden in Registern keine Daten generiert, sondern vorhandene Daten erfasst. Wichtig ist deshalb in diesem Kontext, dass insbesondere geeignete Prozeduren der Erfassung und Codierung sowie der Verarbeitung der Daten Anwendung finden.

Nicht zu vernachlässigen sind als wesentliche Voraussetzungen für eine hohe Datenqualität ausreichendes und qualifiziertes Personal sowie geeignete räumliche und technische Ausstattung.

## VI.2.3 Beschreibung ausgewählter Indikatoren der Leitlinie in Registern

Die TMF-Indikatoren wurden gezielt auf das klinische Register zu den angeborenen Herzfehlern angewandt. Im Folgenden wird beispielhaft dargestellt, wie die Qualitätsindikatoren zum Einsatz kommen (s. Tab. 8).

Tab. 8 Liste von Qualitätsindikatoren mit besonderer Relevanz für Register

| TMF-ID | Qualitätsindikator | Anwendungsbezogener Kommentar |
|---|---|---|
| **Ebene Integrität** | | |
| TMF-1001 | Übereinstimmung mit Vorwert | Neben unveränderlichen Werten wie Geburtsdatum sind über die Zeit Veränderungen von Werten möglich und damit plausibel (z.B. Veränderung der Körpergröße – Wachstum eines Kindes). Deshalb Bildung von Grenzwerten, die sich als nicht plausibel erweisen (Abnahme der Größe mit der Zeit) |

| TMF-ID | Qualitätsindikator | Anwendungsbezogener Kommentar |
|---|---|---|
| TMF-1002 | Konkordanz | Übereinstimmung der in der Datenbank erfassten Daten mit der Quelle; wichtig hierbei ist die Wahl der geeigneten Quelle, da diese als korrekt („wahr“) vorausgesetzt wird. (z.B. Angabe der durchgeführten Operation aus dem Operationsbericht und nicht aus dem Arztbrief).<br>Ebenso Vollständigkeit der Datenerfassung (z.B. alle angeborenen Diagnosen erfasst und auch Ausschlussdiagnostik bestimmter sonst auftretender häufiger Begleitdiagnosen). |
| TMF-1003<br>TMF-1035<br>TMF-1004<br>TMF-1005 | Widerspruchsfreiheit | Festlegungen von Plausibilitäten für sichere Widersprüche (z.B. geschlechtsspezifische genetische Syndrome).<br>Warnung bei möglichen Widersprüchen (z.B. von Art des Herzfehlers abweichende Operationsmethode).<br>Endlosüberlebender (schwierig bei sehr jungen und schwer kranken Patienten) – für sehr gute Datenqualität aufwendige Recherchen über regelmäßige Kontakte und evtl. Einwohnermeldeämter erforderlich. |
| TMF-1013 | Fehlende Einträge/ fehlende Werte bei Datenelementen | Unterscheidung zwischen fehlenden Werten (unvollständige Erfassung von Diagnosen und Behandlungen) und tatsächlich nicht vorhandenen Werten. (Ist in Bezug auf den jeweiligen Patienten zu werten: z.B. pränatale Diagnostik – Eintrag erst ab bestimmten Geburtsjahr relevant, da moderne Diagnostikform, bei Erwachsenen mit AHF irrelevant.) |
| TMF-1016 | Fehlende Einträge/ Datenelemente mit Wert unbekannt | Entscheidung je nach Wert, ob unbekannt akzeptabel oder nicht; z.B. unbekannt akzeptabel bei Angabe zur eigenen Geburt und Schwangerschaft bei erwachsenen Patienten; Jedoch unbekannt möglichst vermeidbar bei Säuglingen. |
| **Ebene Organisation** | | |
| TMF-1028 | Aktualität der gespeicherten Daten | Unterscheidung in unveränderliche (angeborene Erkrankung) und veränderliche Daten (erworbene Erkrankungen, stattgehabte Operationen usw.). Bei veränderlichen Daten gleichzeitige Erfassung des Datums der Datenaktualisierung (Arztbriefdatum, Befunddatum usw.). |
| TMF-1029 | Dubletten | Patientenbezogene Erfassung, keine fallbezogene Erfassung; Dubletten zu vernachlässigen. |
| TMF-1030 | Rekrutierungsrate | Nur sinnvoll für bestimmte Fragestellungen (z.B. Querschnittsuntersuchung wie Erfassung aller Neugeborenen eines Jahrganges oder einer bestimmten Untergruppe mit speziellen Einschluss- und Ausschlusskriterien) – Abgleich mit Geburtenzahlen und Versorgungsdaten<br>Bezug auf Einrichtungen/Zentren – Beachtung des Zuweisungsradius (Einbeziehung des Wohnortes zum Abgleich). |
| TMF-1034 | Vorzeitig aus dem Register ausscheidende Personen | Rücknahme der Einverständniserklärung (sehr selten im Nationalen Register für angeborene Herzfehler, meistens nur bei sehr leichten Herzfehlern), Erfragen der Gründe des Ausscheidens. |




| TMF-ID | Qualitätsindikator | Anwendungsbezogener Kommentar |
|---|---|---|
| TMF-1036<br>TMF-1037 | Synonyme<br>Homonyme | Durch Verwendung des PID-Generators Wahrscheinlichkeit gering, Überprüfung durch Abgleich mit Diagnosen. |
| **Ebene Richtigkeit** | | |
| TMF-1043 | Genauigkeit | Unterschiedliche Verwendung der Nomenklatur für die Diagnose des Herzfehlers/Erkrankung von verschiedenen Ärzten/Kliniken. Vereinheitlichung der Nomenklatur durch Verwendung des IPC-Codes nach Wertung aller Angaben im Arztbrief und Befunden durch besonders geschultes fachspezifisches Personal (Vorteil: hohe Datenqualität bei Phänotypisierung für begleitende Biomaterialbank; Nachteil: sehr personal- und zeitaufwendig). Deshalb Einführung verschiedener Qualitätsstufen für die erforderliche Genauigkeit der Registerdaten. |
| TMF-1044 | Übereinstimmung mit den Originaldaten bezogen auf die Datenelemente | Hoher Anspruch an die Übereinstimmung bei Angaben der Diagnosen von Patienten mit vorhandenem Biomaterial. 4-Augenprinzip, um Qualität zu optimieren. Nachteil: sehr personal- und zeitintensiv. |
| TMF-1045 | Übereinstimmung mit den Originaldaten bezogen auf die Beobachtungseinheiten | Beobachtungseinheit ist ein Patient – Genauigkeit der Übertragung aller Angaben aus Arztbriefen.<br>Hohe Anforderung bei Patienten mit komplizierten Krankheitsverläufen (vielen Interventionen, Komplikationen) |
| TMF-1046 | Vollständigkeit | Bei einem breiten Register mit verschiedenen, teilweise sehr seltenen angeborenen Herzerkrankungen müssen verschiedene Subgruppen unterschiedlich in ihrer Vollständigkeit betrachtet werden, um den Dokumentationsaufwand machbar und bezahlbar zu halten:<br>1. maximale Vollständigkeit der Daten – Patienten mit Biomaterial<br>2. mittlere Vollständigkeit – komplexe, wissenschaftlich sehr interessante Fälle<br>3. variable Vollständigkeit – Patienten mit häufigen und wissenschaftlich weniger interessanten Herzfehlern – Vervollständigung nur nach Vorliegen bestimmter Forschungsfragen<br>Bei epidemiologischem Ansatz: Bezug der Vollständigkeit auf Geburtenraten, Operationszahlen u.a. Angaben aus Erfassungsstatistiken. |
| TMF-1047 | Übereinstimmung mit den Verfahrensregeln | Regelmäßige Aktualisierung der Registerdaten. Klare Verfahrensregeln für verschiedene Untergruppen im Register. |
| TMF-1048 | Repräsentativität | Erfasste Population entspricht Stichprobe: Prozentuale Verteilung der Häufigkeiten der Herzfehler (seltene Fehler – selten im Register; häufige Fehler – häufig im Register). Ebenso Altersverteilung und Geschlechterverteilung – bestimmte Herzfehler haben eine Geschlechterwendigkeit. |

## VI.2.4 Beispielhafte Implementierung von Qualitätsindikatoren im Nationalen Register für angeborene Herzfehler e.V.

Alle Daten werden personenbezogen erhoben, um verschiedene Datensätze eindeutig zuordnen und Einträge zu einzelnen Teilnehmern fortschreiben zu können. Aus datenschutzrechtlichen Gründen arbeitet das Nationale Register für angeborene Herzfehler deshalb mit zwei voneinander getrennten Datenbanken – derjenigen zur Speicherung der personenidentifizierenden Daten (IDAT-DB) und einer weiteren zur Speicherung und Verarbeitung der medizinischen Daten (MDAT-DB).

### VI.2.4.1 Personenidentifizierende Daten (IDAT)

Bei der Eingabe der persönlichen Daten eines Teilnehmers in die IDAT-DB wird anhand der IDAT ein Pseudonym (Patientenidentifikator: PID) erzeugt, das für jede Person eindeutig ist und in allen Datenbanken als Zuordnungskriterium genutzt wird. Phonetische Algorithmen beachten die unterschiedlichen Schreibweisen wie z.B. „Müller" und „Mueller". Der PID bleibt ebenfalls für einen Patienten identisch, wenn er heiratet oder umzieht. So werden einerseits Dubletten vermieden; andererseits ermöglicht diese Art der Pseudonymisierung das kontinuierliche Fortschreiben einzelner Datensätze. Der PID selbst als alphanumerischer Code steht nicht in direktem Zusammenhang mit den zugehörigen IDAT, sodass ein „Rückrechnen" zu den identifizierenden Daten nicht möglich ist. Für die Erzeugung des PID wird ein so genannter externer PID-Generator eingesetzt, auf den die Register-Software verschlüsselt über eine geschützte Verbindung zugreift. Durch die Sensibilität der personenidentifizierenden Daten, die sich in der Patientenliste des PID-Generators befinden, ist dieser Bereich stark geschützt und erlaubt nur einen sehr restriktiven Zugriff.

Der PID-Generator wurde im Kompetenznetz für Pädiatrische Onkologie und Hämatologie (Klaus Pommerening/Universität Mainz) als Werkzeug für die Erzeugung von pseudonymen Patientenidentifikatoren entwickelt (Reng et al. 2006). Er wird in vielen medizinischen Forschungsnetzen eingesetzt und hat sich durch seine hohe Sicherheit und das fehlertolerante Matchingverfahren als Kernmodul für das Identitätsmanagement etabliert. Durch die Eindeutigkeit der PID als Ordnungskriterium in allen Datenbanken des Netzes ist eine sichere und zuverlässige Arbeitsweise mit Sicherstellung des Datenschutzes gewährleistet.

Durch regelmäßiges Kontaktieren der Registerteilnehmer werden die Adressen gepflegt und bei Bedarf über das Einwohnermeldeamt aktualisiert. Dadurch wird auch bekannt, wenn Patienten versterben.

Als sehr effektiv hat sich bei der Datenpflege und Erreichung einer hohen Datenqualität erwiesen, dass die Erzeugung des PIDs anhand der Papiermel-

dung zentral erfolgt und die Patienten immer wieder kontaktiert bzw. auch über einen eigene Webseite informiert werden. So entsteht eine besondere „Kundenbindung“ und die Drop-outs sind sehr gering.

### VI.2.4.2 Medizinische Daten (MDAT)

Bei Eintritt in das Register wird die Krankengeschichte zunächst retrospektiv verfolgt, indem alle angeborenen Erkrankungen und deren Verlauf bis zum Zeitpunkt der Anmeldung dokumentiert werden. Von diesem Zeitpunkt an wird die weitere Krankengeschichte mit eventuell auftretenden erworbenen Erkrankungen, medizinischen Maßnahmen und Eingriffen etc. sodann prospektiv erfasst. Als Datenquelle dienen dabei zum einen subjektive anamnestische Angaben, also Selbstangaben der Teilnehmer; zum anderen werden ärztliche Dokumente wie Arztbriefe, Katheterprotokolle, Operationsberichte oder Autopsiebefunde als Quelle genutzt. Für größtmögliche Genauigkeit und Transparenz wird die Datenquelle (Herkunft, Datum) dabei stets genau dokumentiert.

Die Daten werden ausschließlich zentral durch geschultes Personal in die Datenbank eingegeben. Alle genannten medizinischen Dokumente liegen dem Register-Personal in Papierform vor. Sind die verfügbaren Angaben nicht ausreichend, behält sich das Register die Möglichkeit vor, weitere Dokumente vom betreuenden Arzt/Zentrum anzufordern. Die Genehmigung hierfür wird bereits bei der Anmeldung neuer Teilnehmer mittels *Informed Consent* eingeholt.

Bei der Fülle der gesammelten Daten ist ein etabliertes Ordnungssystem von besonderer Bedeutung, um die Nutzbarkeit der Daten zu gewährleisten. Die Dokumentation der Diagnosen in der Datenbank folgt deshalb einer festgelegten Hierarchie. Dabei wird hinsichtlich der angeborenen Erkrankungen eine Einteilung unter anderem in führende Hauptdiagnosen und Nebendiagnosen vorgenommen. Zudem werden Prozeduren mit Angaben zu Zeit und Ort in chronologischer Reihenfolge erfasst.

Ebenso unabdingbar ist eine einheitliche Codierung der erfassten Daten. Als Nomenklatur nutzt das Nationale Register für alle Diagnosen und Prozeduren den IPC-Code; für erworbene Erkrankungen wird zusätzlich auf den ICD-10-Code zurückgegriffen. Auf Freitext wird weitestgehend verzichtet. Dadurch wird eine standardisierte Klassifizierung der erfassten Diagnosen und Prozeduren gewährleistet, wodurch Homonyme vermieden und Daten schnell und unproblematisch abgefragt bzw. genutzt werden können.

Die Eingabe sämtlicher medizinischer Daten in die Datenbank erfolgt durch geschulte Medizin-Studenten und wird durch einen verantwortlichen Arzt ständig kontrolliert. Liegt zu einem Teilnehmer neben den medizinischen Angaben eine Biomaterial-Probe vor, erfolgt zusätzlich eine zweite Stufe der Qualitätssicherung mittels Supervision. Der jeweilige Qualitätsgrad der Daten wird in der Datenbank vermerkt.

Für alle Arbeitsabläufe innerhalb des Nationalen Registers wurden spezifische *Standard Operating Procedures* (SOPs) entwickelt, welche kontinuierlich an rechtliche, technologische oder infrastrukturelle Änderungen angepasst werden. So sind die folgenden Vorgänge durch SOPs geregelt:

- Nutzung der IT-Plattform
- Organisatorische Fragen wie Zugangsrechte, regelmäßige Anleitung der Mitarbeiter, räumliche und organisatorische Trennung der Datenbanken (IDAT-DB, MDAT-DB)
- Registrierungsprozedur, Eingabe der IDAT, PID-Generierung, Teilnehmer-Feedback
- Beschriftungs- und Ablagesystem von Originaldokumenten
- Handling von originalen Arztbriefen und Eingabe der MDAT
- Qualitätskontrolle der MDAT-Eingabe
- Kontaktaufnahme mit Register-Teilnehmern (leitlinienbasierte Telefoninterviews)
- Abfragen in der MDAT-DB, Zusammenstellen von Daten für geplante Forschungsprojekte
- Herausgabe von Daten für Forschungsprojekte, Generieren einer doppelten Pseudonymisierung

Wie sich zeigt, besteht die besondere Herausforderung in einem Register wie dem Nationalen Register für angeborene Herzfehler in der Phänotypisierung, d.h. der genauen Beschreibung von Diagnosen und Prozeduren. Diese ist die Voraussetzung für die Bildung sinnvoller Kohorten z.B. mit Blick auf angeborene Herzfehler mit gemeinsamer Pathomorphologie, nach Schweregrad (bzw. derzeitigem Leistungszustand) oder, für genetische Untersuchungen, mit Blick auf angeborene Herzfehler mit bestimmten angeborenen Nebenerkrankungen, Syndromen o.ä. Je nach Studienvorhaben können dann für diese Kohorten spezielle Untersuchungsparameter gemäß den Qualitätskriterien für Kohorten erfasst werden.

### VI.2.5 Zusammenfassung und Ausblick

Das Management der Datenqualität in Registern verfolgt zwei wesentliche Ziele:

Zum einen hängen die Möglichkeiten der wissenschaftlichen Nutzung von Daten aus krankheitsbezogenen Registern vor allem von der Qualität der dort erfassten und gespeicherten Daten ab. Je nach Fragestellung der Nutzer können die Anforderungen an die Qualität der Daten sehr unterschiedlich sein. Die Entscheidung über die Verwendbarkeit der Daten kann nur der Nutzer treffen.

Zum anderen lassen sich bei der Betrachtung der Datenqualität wichtige Rückschlüsse auf die etablierten Prozesse der Datenerfassung und Datenspeiche-



rung ziehen. Darauf aufbauend lassen sich geeignete Maßnahmen treffen, um identifizierte Defizitbereiche zu beseitigen und effektiv die Datenqualität positiv zu beeinflussen.

Das Ziel der Sicherstellung einer hohen Datenqualität sollte immer im Zusammenhang mit der Machbarkeit und Finanzierbarkeit betrachtet werden. Die Leitlinien sind ein wichtiger Bezugspunkt, um bereits bei der Planung von Registern sinnvolle Entscheidungen hinsichtlich Datenqualität und -management zu treffen.

## VI.3 Anwendungsempfehlungen für Kohorten

Carsten Oliver Schmidt

### VI.3.1 Einleitung

Kohortenstudien sind in den Lebenswissenschaften ein unverzichtbares Studiendesign zur Beantwortung von Fragestellungen hinsichtlich der Verbreitung, des Verlaufes und Zusammenhangs von Erkrankungen, subklinischen Endpunkten und Risikofaktoren. Eine hohe Datenqualität ist eine notwendige Voraussetzung, um eine hohe Aussagekraft der Studien zu gewährleisten. Einen Orientierungsrahmen bieten die „Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis", kurz GEP [DAE 2004]. Empfehlungen umfassen u.a. eine begleitende Qualitätssicherung aller relevanten Instrumente und Verfahren, eine Dokumentation für Verfahrensänderungen, eine Zweit- und Prüfeingabe für numerische Variablen sowie Plausibilitätskontrollen in einem geeigneten infrastrukturellen Rahmen zur Datenhaltung. In Hinblick auf konkrete Indikatoren zur Datenqualität und deren Operationalisierung bleibt die GEP jedoch unspezifisch. Solche praktisch anwendbaren Indikatoren bietet die vorliegende Fassung der TMF-Leitlinie Datenqualität. Sie wurde gegenüber der letzten Fassung um mehrere Indikatoren erweitert, die im Rahmen der Qualitätssicherung von Beobachtungsstudien relevant sind.

Der folgende Überblick erläutert spezifische Aspekte der Sicherung von Datenqualität in Kohortenstudien, führt ausgewählte Indikatoren der TMF-Leitlinie Datenqualität ein und beschreibt am Beispiel der Study of Health in Pomerania, SHIP [Völzke et al. 2011], eine exemplarische Umsetzung im Rahmen eines kontinuierlichen Datenmonitorings. Die SHIP-Studie ist eine bevölkerungsbezogene Gesundheitsstudie, bestehend aus zwei Kohorten mit insgesamt über 8800 Teilnehmern, die im Rahmen des Forschungsverbundes „Community Medicine" an der Universitätsmedizin Greifswald durchgeführt wird. Die Studie erfasst erstens die Prävalenz und Inzidenz von Erkrankungen, subklinischen Auffälligkeiten und Risikofaktoren. Zweitens wird der Zusammenhang zwischen diesen im Quer- und Längsschnitt untersucht.

Insofern Kohortenstudien auch die Nutzung von sekundären Datenquellen vorsehen, etwa die Nutzung von Daten aus klinischen Tumorregistern, Krankenhäusern oder der kassenärztlichen Vereinigung, greifen primär Indikatoren zur Qualitätskontrolle bei Registern (s. Kap. VI.2), auf die hier nicht eingegangen wird. Umgekehrt bieten sich wichtige Möglichkeiten zur Validierung von Sekundärdaten auf Basis von Primärdaten aus Kohortenstudien, was im Ausblick beschrieben wird.

### VI.3.2 Schwerpunkte der Qualitätssicherung in Kohortenstudien

Im Gegensatz zu Registern und Datenrepositorien fokussiert die Qualitätssicherung in Kohortenstudien auf den Prozess der Datengenerierung. Daten sollen reliabel, valide und vollständig erhoben werden. Hierin besteht eine wesentliche Gemeinsamkeit mit klinischen Studien, wobei in Hinblick auf letztere ein wesentlich engerer, in Teilen gesetzlich geregelter Rahmen zur Qualitätssicherung besteht [Baigent et al. 2008, Eberhard/Herrlinger 2010, Macefield et al. 2013].

Primäre Ziele der Qualitätssicherung in Kohortenstudien sind die Minimierung von Messfehlern, die Gewährleistung einer hohen Generalisierbarkeit sowie der internen Vergleichbarkeit der Messungen innerhalb und zwischen den Erhebungswellen. Dazu sind zahlreiche Maßnahmen auf konzeptueller, struktureller und prozessualer Ebene notwendig [DAE 2004, Rajaraman/Samet 2005, Whitney et al. 1998, Lüdemann et al. 2000].

Konzeptuell ist ein geeignetes Studienprotokoll grundlegend. Dieses definiert u.a. Fragestellungen und Hypothesen, regelt die Stichprobenauswahl, Rekrutierungsmethoden, die Instrumentenauswahl und andere methodische Fragen. Viele der im Folgeabschnitt genannten Indikatoren liefern nur unter Verwendung standardisierter und vorab validierter Messverfahren sinnvolle Ergebnisse.

Hinsichtlich der strukturellen Ebene sind wesentliche Voraussetzungen zur Gewährleistung einer hohen Datenqualität u.a. geeignete Räumlichkeiten, die Qualifikation des Personals, Organisationsstrukturen mit einer nach Möglichkeit personellen Trennung von Untersucher- und Qualitätssicherungsfunktionalitäten, die IT-Infrastruktur zur Dateneingabe, u.a. elektronische Datenerfassungssysteme (Electronic Data Capture, EDC), Datenspeicherung und Datenmanagement.

Prozessual kann zwischen Maßnahmen unterschieden werden, die im Vorfeld der Datenerhebung stattfinden und solchen, die an den generierten Daten während und nach der Dateneingabe ansetzen. Ersteres umfasst u.a. Maßnahmen zur Qualifizierung des Personals, die Sicherstellung einer geeigneten Gerätefunktionalität und letzteres Maßnahmen zur Eingabe und zum Monitoring der generierten Daten. Einen Überblick wichtiger Maßnahmen bietet Tabelle 9.




Tab. 9 Prozessuale Elemente der Qualitätssicherung in Kohortenstudien

| Maßnahmen vor der Datenerhebung |
|---|
| Studienhandbücher für Untersuchungen, Datenmanagement, Qualitätssicherung |
| Checklisten zur Untersuchungsdurchführung |
| Training/(Re-)Zertifizierung von Untersuchern, Befundern und ggfs. Trainern bei multizentrischen Projekten |
| Gerätekalibrierung, Gerätevergleiche, Gerätewartung |
| Phantommessungen bei bildgebenden Verfahren |
| Gespräche/Qualitätszirkel zu qualitätsrelevanten Aspekten der Untersuchung |
| Pilotierung der Studie/Prätest einzelner Untersuchungsmodule |
| **Maßnahmen während und nach der Datenerhebung** |
| Kontrolle der Teilnahmevoraussetzungen (Identität und Einverständnisse) |
| Standardisierte Dateneingaben (z.B. webbasiert mit automatischer Plausibilitätskontrolle) |
| Mehrfachbefundungen (u.a. bei bildgebenden Verfahren) |
| Mehrfacheingaben (Standard bei Eingabe von Papiervorlagen, z.B. Fragebögen) |
| On-Site Monitoring (Räumlichkeiten sowie Untersuchungsprozess) |
| Standardisierte und zentralisierte Erfassung von Auffälligkeiten (z.B. datenbankbasiert) |
| Qualitätsberichte |
| Monitoring der laufenden Datenerhebung (Fokus auf Rohdaten) |
| Datenmanagement, syntaxbasierte Datenbereinigungen |
| Externes Monitoring (Advisory Board) |

Zur Kontrolle der laufenden Datenerhebungen ist ein zentrales Datenmonitoring, wie es auch in die auch in randomisierten Studien empfohlen ist [Baigent et al. 2008, De 2011], wesentlich, um Datenauffälligkeiten effizient detektieren zu können.

Die Implementation von Maßnahmen zur Qualitätssicherung und Qualitätskontrolle ist als Rückkopplungsschleife zu verstehen, bei der alle Maßnahmen während und nach der Datenerhebung alle Maßnahmen bedingen können, die vor der Datenerhebung genannt sind. Dies schließt auch Inhalte der konzeptuellen und strukturellen Ebene ein, die sowohl vor als auch während der Studie Gegenstand von Anpassungen sein können. Insgesamt entspricht dies dem in der Einleitung genannten Anspruch auf eine kontinuierliche Verbesserung der Datenqualität (s. Abb. 1).

Nach Abschluss der Datenerhebung folgt die finale Datenbereinigung, die vollständig skriptbasiert zu erfolgen hat, um alle Unterschiede zwischen Roh- und Auswertungsdaten nachvollziehen zu können. Anschließend greifen wei-

tere qualitätssichernde Prozesse zum geeigneten Umgang mit den bereinigten Ergebnisdaten. Beispiele sind ein standardisiertes, skriptbasiertes Datenübergabeprozedere oder ein Publikationskomitee. Durch einen Zugriff auf Auswertungssyntax und Manuskripte erhält Letzteres, wie z.B. in der SHIP-Studie, vor der Publikation die Möglichkeit, Autoren Hinweise zu geben.

Eine besondere Herausforderung in komplexen Kohortenstudien wie SHIP, der Nationalen Kohorte [Wichmann et al. 2012] oder dem WHO-MONICA Projekt [Tolonen et al. 2006] besteht darin, die Vergleichbarkeit der Messwerte zwischen verschiedenen Untersuchern, Zentren, Labors, Geräten über einen langen Zeitraum zu gewährleisten. Zentral ist also die interne Validität der Messergebnisse innerhalb einer Studie. Daher ist es nicht das primäre Ziel, die neuesten Gerätegenerationen oder Softwareupdates umgehend zu implementieren, sondern eine einmal gewählte Technologie so lange wie möglich beizubehalten und jeden Wechsel zwischen Technologien und Methoden durch Vergleichsuntersuchungen zu flankieren. Relativ kurze Lebenszyklen technischer Geräten erschweren dies, wenn alte Geräte nicht mehr zur Verfügung stehen. Durch eine erhöhte Sensitivität z.B. moderner bildgebender Technologien werden morphologisch ähnliche Endpunkte anders visualisiert, Detektionsraten klinischer Endpunkte verändert und dadurch die Vergleichbarkeit zwischen den Erhebungswellen gesenkt. Ein weiteres Beispiel sind zwischen und innerhalb von Erhebungswellen auftretende Veränderungen der laboranalytischen Untersuchungsmethoden, was kritisch ist, da die Variabilität der Messwerte zwischen Assays erheblich sein kann [Lai et al. 2012].

Ein Datenaudit, bei dem ein externes Team die Forschungsdaten mit den Prüfbögen vergleicht, spielt im Vergleich zu klinischen Studien aufgrund der anderen rechtlichen und finanziellen Rahmenbedingungen eine untergeordnete Rolle. Zudem wird durch ein solches Vorgehen nur ein Teil qualitätsrelevanter Fehler detektiert [Nahm et al. 2008]. Nichtsdestotrotz ist eine Quelldatenverifizierung auch in Kohortenstudien in Bezug auf ausgewählte Untersuchungen Routine. Ein Beispiel ist der Abgleich von Papiervorlagen mit den elektronischen Daten, etwa bei Fragebögen, Einverständniserklärungen oder Geräten, die keine automatische Übertragung von Messwerten erlauben [Lüdemann et al. 2000].

### VI.3.3 Beschreibung ausgewählter Indikatoren der Leitlinie in Kohortenstudien

Die im Rahmen dieser TMF-Leitlinie beschriebenen Indikatoren der Datenqualität fokussieren auf die Ergebnisqualität vorliegender Daten. Nicht im Fokus stehen also die beschriebenen konzeptuellen und strukturellen Rahmenbedingungen sowie qualitätssichernde Maßnahmen vor und während der Datenerhebung. Dies war mit der vorliegenden Version der Leitlinie auch nicht intendiert. Damit spielen die TMF-Indikatoren vor allem im Datenmonitoring eine Rolle.

Methodisch lassen sich die drei Ebenen der TMF-Leitlinie vereinfacht wie folgt charakterisieren: Die Ebene Integrität fokussiert auf die Reliabilität der Messwerte und das Fehlen einzelner Datenelemente (Vollständigkeit) sowie im geringeren Maße auf die interne Validität. Die Ebene Organisation fokussiert auf Indikatoren für das Fehlen kompletter Beobachtungseinheiten (Vollzähligkeit), was für die Generalisierbarkeit wesentlich ist, und die Ebene Richtigkeit auf die externe Validität. Allerdings bestehen zwischen den Indikatoren der drei Ebenen der TMF-Leitlinie teils enge Bezüge, weshalb die eindeutige Zuordnung nur begrenzt möglich ist.

Gegenüber der alten TMF-Leitlinie Datenqualität [Nonnemacher et al. 2007] bestehen Veränderungen mit Relevanz für Kohortenstudien vor allem in der Verfeinerung und Ergänzung von Indikatoren auf der Ebene Integrität [Stausberg et al. 2012]. Ein Beispiel ist die Differenzierung verschiedener Ursachen von Messfehlern. Eine kommentierte Übersicht ausgewählter Indikatoren mit Relevanz für Kohortenstudien wird in Tabelle 10 gegeben.

Vergleicht man die drei Ebenen der TMF-Qualitätsindikatoren zwischen Registern und Kohortenstudien, zeigt sich deren unterschiedliche Gewichtung in der praktischen Anwendung. Diese ist in unterschiedlichen Vorgehensweisen in beiden Anwendungsfeldern begründet, Qualitätsindikatoren zu bilden. Kohortenstudie bieten bessere Möglichkeiten, Indikatoren zur Reliabilität und fehlenden Werten zu generieren als zur internen oder externen Validität. Meistens fehlt eine externe Quelle, die zur Überprüfung der Validität der Daten dienen kann. Bei Registern besteht hingegen ein starker Fokus auf die Ebene Richtigkeit im Sinne einer Quelldatenverifizierung. Sofern dies in Kohortenstudien nicht möglich ist, stellt dies keine qualitative Überlegenheit der Register dar. Im Gegenteil: Inwieweit die einem Register zu Grunde liegenden Quellen ihre Daten reliabel und valide erfassen, bleibt weitgehend offen, während genau dies in Kohortenstudien zentraler Gegenstand der Qualitätskontrolle ist und durch geeignete Qualitätssicherungsmaßnahmen gewährleistet werden kann.

Bei der Anwendung von Schwellenwerten der Leitlinie sind im Rahmen der Kohortenstudie zwei Ebenen zu unterscheiden: Erstens Schwellen zur Entscheidung über Maßnahmen der Qualitätssicherung im Rahmen einer laufenden Datenerhebung. Diese fokussieren auf Rohdaten, eine Implementierung wird beispielhaft im folgenden Abschnitt beschrieben. Zweitens Schwellen zur abschließenden Bewertung der erreichten Datenqualität auf der Ebene des Auswertungsdatensatzes. Letzteres kann für die Planung von Datenanalysen genutzt werden, etwa durch Stratifizierung der Analysen über verschiedene Qualitätsstufen bei multizentrischen Studien [Tolonen, et al. 2006]. Ein weiteres Anwendungsgebiet ist die Nutzung der Indikatoren zur retrospektiven Harmonisierung vorliegender Daten, um die Qualität von Analysen aus gepoolten Studien zu verbessern [Fortier et al. 2011].

Tab. 10 Liste von Qualitätsindikatoren mit besonderer Relevanz für Kohortenstudien

| Qualitätsindikator | Anwendungsbezogener Kommentar |
|---|---|
| Übereinstimmung mit Vorwert (TMF-1001) | Die Bildung von Grenzwerten für Werte, die sich unplausibel von ihren Vorwerten unterscheiden, ist zur Beurteilung der internen Validität wichtig. Da intraindividuell oft große Veränderungen über die Zeit plausibel sind, kommt Lagemaßen (z.B. Mittelwert) für ein Untersuchungskollektiv eine wichtige Rolle zu. |
| **Werteverteilung (TMF-1006)** | |
| Bevorzugung bestimmter Endziffern (TMF-1007) | Anfällig sind Messungen mit nicht-digitaler Anzeige, bei denen manuell gerundet werden muss. Entsprechend wird das Auftreten durch geeignete Gerätewahl mit möglichst automatischer Datenübertragung minimiert. |
| Werteverteilung der durch Untersucher erfassten Parameter (TMF-1009) | Hauptursache für Messfehler in Kohorten, der Messfehler schwankt stark zwischen Untersuchungstypen. Eine randomisierte Zuweisung zwischen Untersuchern und Probanden ist wichtig, damit im Durchschnitt keine Messunterschiede zwischen Untersuchern auftreten. Darauf beruht die Bildung statistischer Kennwerte. |
| Werteverteilung der durch Geräte erfassten Parameter (TMF-1010) | Wird in der Praxis auch bei geeichten Geräten regelmäßig beobachtet und bedarf fortwährender Kontrolle. Eine randomisierte Zuweisung, sowohl gegenüber Probanden als auch Untersuchern ist für die Bildung von statistischen Kennwerten wichtig. |
| Werteverteilung von Befunden (TMF-1011) | Quantitative Befundungen sind im Allgemeinen weniger anfällig für relevante systematische Fehler als qualitative. Insbesondere die qualitative Beurteilung bildgebender Ergebnisse erweist sich in der praktischen Umsetzung als problematisch. |
| Werteverteilung von Parametern zwischen Zentren (TMF-1052) | Ist eine Variante der TMF-Indikatoren 1009 und 1010, bei der mittlere Unterschiede zwischen Untersuchungszentren fokussiert werden. Die Interpretation ist dadurch erschwert, dass Populationsunterschiede mit Messfehlern konfundiert sind, dem nur in Teilen durch eine adjustierte Analyse begegnet werden kann. |
| **Fehlende Einträge** | |
| Fehlende Module (TMF-1012) | Bezeichnet die Nichtrealisierung von kompletten Untersuchungsmodulen und kann sowohl designbedingt sein als auch durch Selbstselektion. In Hinblick auf letzteres siehe TMF-1032. |
| Fehlende Werte bei Datenelementen (TMF-1013) | Siehe 1014 |
| Fehlende Werte bei mandatorischen Datenelementen (TMF-1014) | Fast alle Datenelemente in Kohortenstudien sind mandatorisch. Sofern geeignete Eingabekontrollen vorliegen, etwa durch EDC mit Plausibilitätskontrollen, sind fehlende Werte mit 0% Toleranzschwelle zu bearbeiten. Kritischer sind Fragebögen, da Probanden bei bestimmten Fragen mit erhöhter Wahrscheinlichkeit keine Angaben machen. |




| Qualitätsindikator | Anwendungsbezogener Kommentar |
|---|---|
| Fehlende Werte bei optionalen Datenelementen (TMF-1015) | Spielen in Kohortenstudien eine untergeordnete Rolle. Sind eher bei Nutzung von Daten aus der klinischen Routine relevant. |
| Datenelemente mit Wert unbekannt o.ä. (TMF-1016) | Sollten bei elektronischen Eingaben in geschlossenen Ergebnisformaten nicht aufgetreten, daher 0% Toleranz |
| Ausreißer bei stetigen Datenelementen (TMF-1018) | Jeder Ausreißer bei vordefinierten Plausibilitätsgrenzen ist Gegenstand von Qualitätssicherungsmaßnahmen, 0% Toleranz. |
| Werte, die die Messbarkeitsgrenzen von Verfahren unter- oder überschreiten (TMF-1019) | Im Kontext von Laboranalysen bei Bevölkerungskohorten relevant, da durch den hohen Anteil gesunder Probanden Messbarkeitsgrenzen systematisch unterschritten sein können. |
| **Unerlaubte Werte** | |
| Unerlaubte Werte bei qualitativen Datenelementen (TMF-1021) | Sollten bei EDC in geschlossenen Ergebnisformaten nicht auftreten, daher 0% Toleranz. Jeder Fall ist Gegenstand von Qualitätssicherungsmaßnahmen. |
| Unerlaubte Werte bei qualitativen Datenelementen zur Kodierung von Missings (TMF-1022) | Sollten bei EDC in geschlossenen Ergebnisformaten nicht auftreten, daher 0% Toleranz. Jeder Fall ist Gegenstand von Qualitätssicherungsmaßnahmen. |
| Unerlaubte Werte zur Kodierung von fehlenden Modulen (TMF-1023) | Sollten durch Definition von geeigneten Sprungregeln nicht auftreten, daher 0% Toleranz. Jeder Fall ist Gegenstand von Qualitätssicherungsmaßnahmen. |
| Unerlaubte Werte bei quantitativen Datenelementen zur Kodierung von Unter- oder Überschreitungen von Messbarkeitsgrenzen (TMF-1024) | Sollten bei EDC in geschlossenen Ergebnisformaten oder bei automatischem Übertrag aus Messgeräten nicht auftreten, daher 0% Toleranz. Jeder Fall ist Gegenstand von Qualitätssicherungsmaßnahmen. |
| Nachweis bekannter Korrelationen (TMF-1027) | Kann als Element der Qualitätskontrolle unter Nutzung von verfügbaren Informationen aus Vorwellen oder anderen Studien mit ähnlicher Methodik genutzt werden. Problematisch ist die Definition von zulässigen Grenzen für eine Bewertung qualitätsrelevanter Abweichungen. |

### VI.3.4 Beispielhafte Implementierung von Qualitätsindikatoren in der SHIP-Studie

SHIP ist nicht auf die Erfassung bestimmter vordefinierter Erkrankungen, subklinischer Auffälligkeiten oder Risikofaktoren ausgelegt [Völzke et al. 2011]. Vielmehr werden gesundheitsrelevante Endpunkte in einer maximal möglichen Breite gemessen. Entsprechend umfasste SHIP in der letzten Erhebungswelle mehrere dutzend Untersuchungen mit einer Gesamtdauer von über 20 Stunden. Diese reichten von einem Interview und Fragebögen über eine Blutdruckmessung, kardiovaskuläre Untersuchungen, bildgebende Verfahren incl. Ultraschall und MRT bis hin zu einer dermatologischen Untersuchung und Schlaflabor. Dadurch entstehen sehr hohe Anforderungen an die Imple-

mentation einer leistungsfähigen Qualitätssicherung, im Rahmen derer ein begrenzter Mitarbeiterstab vielfältige Untersuchungen mit heterogenen Anforderungen betreuen muss.

Strukturell ist das Organisationszentrum für die Studienplanung verantwortlich, während das Untersuchungszentrum die SHIP-Kernuntersuchungen durchführt. Hinzu kommen assoziierte Kliniken, die spezielle Untersuchungsmodule durchführen, z.B. die Radiologie oder Dermatologie. Für jede Untersuchung sind auf Seiten des Untersuchungszentrums Untersuchungs- und Geräteverantwortliche definiert, auf Seiten des Organisationszentrums ein Qualitätsverantwortlicher. Dieser ist für die geeignete Implementierung und Durchführung der Untersuchungen verantwortlich.

Soweit das Monitoring von Daten im Vordergrund steht, kommt teilautomatisch generierten Qualitätsreports bei SHIP eine herausgehobene Rolle zu. Diese werden für alle Untersuchungsmodule generiert und operationalisieren zahlreiche Indikatoren der TMF-Leitlinie. Die vierteljährlich, bei Bedarf auch öfter generierten Qualitätsreports haben folgende Merkmale und Voraussetzungen:

- In den verschiedensten Untersuchungsmodulen sind ähnliche Indikatoren sinnvoll einsetzbar. Im Fokus der Reports stehen die Anzahl fehlender Werte (TMF-1012–1016), Extremwerte (TMF-1018), die Variabilität von Messwerte zwischen Untersuchern (TMF-1009), Geräten (TMF-1010) und Befundern (TMF 1011), die Bevorzugung bestimmter Endziffern (TMF-1007), sowie Zeittrends (TMF-1009–1011). Unerlaubte Werte (TMF-1021–1026) werden bereits vorab im Datenmanagement erfasst.
- Hauptzielsetzung der Berichte ist die kontinuierliche Verbesserung der Datenqualität während der laufenden Studie. Daher ergeben sich strenge Grenzwerte für die Einleitung qualitätssichernder Maßnahmen, die bereits bei Verdacht auf Probleme ansetzen.
- Der statistische Berichtsteil wird inkl. der Tabellen und Grafiken durch Skripte automatisch generiert, manuell erfolgt lediglich die Interpretation durch die Qualitätsverantwortlichen, die in einer Datenbank über eine Webeingabemaske eingetragen wird. Die automatische Zusammenstellung ist Voraussetzung zur breiten Abdeckung der erhobenen Variablen. Der Report wird über LaTeX erstellt und als PDF-Dokument ausgegeben.
- Alle erhobenen Variablen unterliegen einem Screening auf Auffälligkeiten, dies umfasst die o.g. Aspekte exklusive Endziffern und Zeittrends. Ein Teil der Variablen wird mit Fokus auf die Variabilität der Messwerte und Zeittrends vertieft analysiert und graphisch dargestellt. Über die Zuordnung zu einer vertieften Analyse entscheiden die Bedeutung einer Variable für spätere Analyseprojekte und das erwartete Risiko für Datenauffälligkeiten.
- Die Zuweisung von Probanden zu Untersuchern erfolgt annähernd randomisiert, da die Terminvergabe Probandenmerkmale nicht berücksich-




tigt. Aus dieser quasi-experimentellen Situation ergibt sich, dass der Erwartungswert für beobachtete Probandenmerkmale zwischen Untersuchern identisch sein sollte. Abweichungen von Messwerten zwischen Untersuchern weisen unter der Voraussetzung ausreichender Fallzahlen auf systematische Fehler hin, die einer Abklärung bedürfen. Hinsichtlich der Annahme gleicher Erwartungswerte besteht eine wichtige Einschränkung: Probandenmerkmale verändern sich im Verlauf einer Erhebungswelle, ein wichtiger Einflussfaktor sind Rekrutierungsmethoden [Haring et al. 2009]. Daher gilt die Annahme gleicher Erwartungswerte nicht für Untersucher, die zu verschiedenen Zeiträumen tätig sind.

- Alle Indikatoren werden hinsichtlich der Entscheidung über qualitätssichernde Maßnahmen unabhängig voneinander betrachtet. Daher wurde bislang kein übergreifender Score entwickelt und genutzt.

Weitere qualitätsrelevante Indikatoren werden außerhalb der Qualitätsreports gebildet, da sie für keine einzelne Untersuchung spezifisch sind. Dazu zählen die Rekrutierungsrate (TMF 1030), Dubletten (TMF-1029), Verweigerungsraten (TMF-1031) oder vorzeitig ausscheidende Beobachtungseinheiten (TMF-1034).

Am Beispiel der Anwendung von Schwellen zur Entscheidung über Maßnahmen der Qualitätssicherung im Rahmen einer laufenden Datenerhebung in SHIP zeigen sich sehr deutliche Unterschiede zu Registern. Laut TMF-Leitlinie liegen empfohlene Schwellen für den Anteil von Werten mit Auffälligkeiten bei fehlenden Werten bei mandatorischen Datenelementen bei 5%, der Grenzwert für Auffälligkeiten Widerspruchsfreiheit liegt bei 5% und der für Extremwerte bei 10%. Im Gegensatz hierzu liegen die entsprechenden Grenzwerte bei SHIP für diese Felder bei 0%. Der Grund ist, dass aufgrund des EDC weder fehlende noch unbekannte Werte im Rahmen der Datenerhebung auftreten dürfen. Jeder Einzelfall führt zur Überprüfung, um eine Wiederholung zu vermeiden. Auch bei Extremwerten führen vordefinierte Plausibilitätsgrenzen zu einem Routinecheck aller außerhalb liegender Werte mit den Untersuchern und ggfs. den Fachverantwortlichen für die Untersuchung.

Komplizierter gestaltet sich die Bewertung auffälliger Werteverteilungen. Die Variabilität der Messwerte zwischen Untersuchern, Befundern und Geräten wird über Intraklassenkorrelationen (ICC) und Design-Effekte bestimmt sowie im Verlauf graphisch auf Basis von nichtparametrischen Regressionsmodellen visualisiert. Weil verschiedene Untersucher unterschiedliche Probanden untersuchen, indiziert eine niedrige Intraklassenkorrelation wenig systematische Abweichungen. Da sowohl graphische als auch statistische Ergebnisse genutzt werden, ist kein strikter Grenzwert vorgegeben. Üblicherweise werden Variablen mit ICC > 0,02 Gegenstand weiterer Betrachtungen.

Zusätzliche Grenzwerte spielen bei der Auswertung von Zertifizierungsergebnissen eine Rolle, auf deren Basis sich sowohl Maße zur Inter- wie Intra-Rater-Reliabilität bilden lassen. Visualisiert werden Zertifizierungsergebnisse im

Falle kontinuierlicher Variablen über Bland und Altman Plots [Bland/Altman 1999]. Der Standard in SHIP für das Bestehen einer Zertifizierung ist u.a. weniger als 5% durchschnittliche Abweichung (mean bias) im Vergleich zum Goldstandard [Völzke et al. 2011]. Bei dichotomen Outcomes sollte ein Kappa-Wert > 0,8 erreicht sein.

Die erhebliche Komplexität der Implementierung der Studie zur Gewährleistung eines aussagekräftigen Monitorings kann am Beispiel bildgebender Verfahren demonstriert werden. Mögliche Einflussfaktoren für die Ergebnisse sind

1. das Messgerät zur Erstellung der Bilder,
2. der Untersucher, der diese Bilder aufnimmt,
3. der Befunder, der die Bilder auswertet und
4. die Technik, die der Befunder verwendet (z.B. Monitor, Software).

Genaugenommen sind auch die räumlichen Verhältnisse, z.B. angemessene Verdunkelung, unter denen die Bilder gemacht und befundet werden, relevant. Alle Ebenen voneinander zu trennen ist statistisch komplex und kann im Falle von Studien mit mehreren Untersuchern, Geräten und Befundern nur dadurch adäquat gelöst werden, dass erstens eine vollständige Dokumentation dieser Faktoren erfolgt und zweitens diese Einflussfaktoren unabhängig voneinander sind. Verhindert werden muss also, dass z.B. Mitarbeiter A bevorzugt Raum 2 mit Gerät 2 bedient und dies in der Frühschicht, während Mitarbeiter B Raum 1 mit Gerät 1 nimmt und vorwiegend Wochenenden und Nachmittage betreut und beide dazu ihre eigenen Bilder befunden. In diesem Fall ist die Beurteilung der Fehlerquellen untrennbar konfundiert. Letztlich sind also aussagekräftige Metadaten erforderlich, um die Belastbarkeit von statistischen Routinen und Ergebnissen besser deuten zu können, was im Indikator TMF-1050 abgebildet werden kann.

### VI.3.5 Zusammenfassung und Ausblick

Indikatoren zur Messung der Datenqualität in Kohortenstudien sind unverzichtbar und sind in der aktuellen TMF-Leitlinie Datenqualität aufgeführt. Dass diese Indikatoren nicht alle Aspekte der Qualitätssicherung in Kohortenstudien umfassen, liegt am Fokus auf die Bewertung vorliegender Daten. Das Implementationsbeispiel in SHIP verdeutlicht, dass eine praktikable Bildung von Indikatoren in einer komplexen Kohortenstudie IT-Lösungen erfordert. Ein offener Punkt ist die Aggregation verschiedener Indikatoren zu globalen Kennwerten der Studienqualität. Diese Aggregation gestaltet sich kompliziert, da Kohortenstudien ein sehr breites Spektrum von Endpunkten und Methoden umfassen. Nur wenige Indikatoren sind für eine Kohortenstudie von globaler Bedeutung, z.B. die Rekrutierungsrate, während die meisten anderen spezifisch für einzelne Untersuchungsmodule und darin enthaltene Variablengruppen sind. Mit der gelungenen Implementation eines Interviews wird nichts über die Qualität bildgebender Verfahren in der gleichen Studie ausgesagt, da




sich Techniken und Personal unterscheiden. Daher ist z.B. die Bildung von Mittelwerten der Indikatoren über alle Untersuchungsmodule wenig aussagekräftig. Zielführender ist eine Aggregation auf der Ebene einzelner Untersuchungsmodule. Daraus lassen sich Hinweise zu potenziellen Problematiken bei konkreten Analyseprojekten gewinnen.

Geeignete Studiendesigns sind eine wesentliche Voraussetzung, um aussagekräftige Indikatoren zur Datenqualität zu generieren. Einige Designelemente sollten dabei eine größere praktische Bedeutung erhalten als bisher. Beispielsweise würden Validierungsstudien innerhalb einer Kohortenstudie vielfältige Optionen eröffnen, Indikatoren der Ebene Richtigkeit zu generieren. Sofern Sekundärdatenquellen in Kohortenstudien genutzt werden, bieten sich leistungsfähige Möglichkeiten zur Bildung von Indikatoren der Ebene Richtigkeit für die Register selbst. Standardisierte und zumeist besser qualitätsgesicherte Messungen der Kohortenstudien könnten zur Validierung von Angaben aus Sekundärdatenquellen dienen, beispielsweise in Hinblick auf chronische Erkrankungen. Ein Beispiel ist die Abbildung von Hypertonie von Patienten in Daten der Kassenärztlichen Vereinigung im Vergleich zu den standardisierten Messungen einer Kohortenstudie.

## VI.4 Anwendungsempfehlungen für Data Repositories

Thomas Schrader

### VI.4.1 Einleitung

Unter einem klinischen Forschungsdatenregister wird ein Repository verstanden, welches klinische Daten aus unterschiedlichen Quellen für Forschungszwecke zugänglich macht [Eminaga et al. 2013, Hruby et al. 2013, Piwowar/Vision 2013, Sinaci et al. 2013]. Dabei können unterschiedliche Strategien umgesetzt werden:

1. Die Datenhaltung erfolgt in einem zentralen Register als sog. Centralized Research Data Repository [Hruby et al. 2013].
2. Es wird ein föderierter Ansatz verfolgt, der die eigentlichen Daten in den Quellen hält. Die Abfrage erfolgt via Metadaten über die Einzelquellen [Ngouongo et al. 2013, Sinaci et al. 2013, Weber 2013].

Allen gemeinsam ist die Tatsache, dass der Aufbau und die Datenhaltung an keine spezifische Forschungsfrage gebunden sind. Vielmehr steht die Erschließung von klinischen Daten im Vordergrund unter dem Aspekt der Wiederverwendung dieser Daten (secundary use of clinical data).

Das Open European Nephrology Science Center (OpEN.SC) ist ein Forschungsdaten-Repository für klinische Daten aus dem Bereich der Nephrologie. Es war ein von der Deutschen Forschungsgemeinschaft gefördertes Projekt aus dem

Bereich der Digitalen Bibliotheken, welches 2013 abgeschlossen wurde. Entstanden ist eine internetbasierte Plattform mit einer serviceorientierten Architektur (SOA), die in der Lage ist, innerhalb von kurzer Zeit klinische Daten aus unterschiedlichen Einrichtungen und dementsprechend verschiedener Infrastruktur in einem Repository zusammenzuführen und für Forschungsfragen zugänglich zu machen. Aus drei nephrologischen Kliniken sowie dem Institut für Pathologie der Charité, Universitätsmedizin Berlin, wurden ca. 20.000 Fälle für die Forscher erschlossen und zugänglich gemacht. Dabei wurden Krankheitsverläufe von mehr als 15 Jahren ebenso eingeschlossen wie auch vollständig digitalisierte histologische Schnitte als sogenannte Whole Slide Images (WSI).

OpEN.SC verstand sich als eine Einrichtung der Datenhaltung und des Datenmanagements. Die Aufgabe ist mit der einer „Research Data Stewardship" zu vergleichen [Marco 2000]. Aus dieser Rolle im Lebenszyklus wissenschaftlicher Daten ergaben sich zum einen eine besondere Verantwortung und zum anderen eine besondere Sichtweise auf Daten und Datenqualität.

### VI.4.2 Schwerpunkte der Qualitätssicherung in einem Daten-Repository

2007 wurde die vorliegende Leitlinie in Version 1.0 publiziert [Nonnemacher et al. 2007]. Darin wurden zum ersten Mal systematisch für den medizinischen Bereich Datenqualitätskriterien vorgestellt und in den Kontext von Tumorregistern und Kohortenstudien gebracht. 2012 wurden einige Qualitätsparameter ergänzend aufgenommen, die sich als bedeutungsvoll für andere Arten von Datenbeständen wie Kohorten (Study of Health in Pomerania – SHIP, Greifswald) und Forschungsdatenregistern wie OpEN.SC erwiesen haben [Stausberg et al. 2012].

### VI.4.3 Beschreibung ausgewählter Indikatoren der Leitlinie für Forschungsdaten-Repositorien

Im Rahmen des OpEN.SC Projektes wurde sehr frühzeitig begonnen, das Thema Datenqualität aufzugreifen. Dabei wurde in der Diskussion festgestellt, dass die Meinungen und das Verständnis von Qualität sehr stark divergierten. Es wurde klar, dass klinisch tätige Kolleginnen und Kollegen andere Anforderungen an die Daten haben als beispielsweise Pathologinnen und Pathologen. Aus der Sicht des Systemadministrators wurden wiederum andere Aspekte betrachtet. Diese Unterschiede lassen sich dadurch erklären, dass die Daten inhaltlich anders strukturiert sein müssen, wenn eine Verlaufsuntersuchung retrospektiv durchgeführt werden soll oder wenn nach Mustern und Zusammenhängen in Daten gesucht wird.

Für ein Forschungsdaten-Repository ist kennzeichnend, dass die Daten zunächst ohne konkrete Aufgabenstellung gesammelt werden. Die klinischen




und Bild-Daten werden aus unterschiedlichen Quellen, in unterschiedlichem Umfang und mit unterschiedlichem Format im Register abgelegt. Die Analyse von Datenqualität verfolgt deshalb zwei Ziele:

1. Anwender und Nutzer der Daten sollen einen Überblick erhalten, mit welchen Daten und Eigenschaften der Datensammlung sie es zu tun haben. Dazu können Parameter verwendet werden, die das Repository beschreiben.
2. Die Frage nach der Datenqualität soll vor dem Hintergrund einer spezifischen Aufgabenstellung (z.B. Studienrekrutierung) untersucht werden. Zentrale Frage dabei ist, ob die Daten hinsichtlich Inhalt und Form hinreichend gut sind, um verlässliche Aussagen zu erhalten. Dabei können für einige Parameter konkrete Anforderungen formuliert und damit auch bewertet werden.

Um ein allgemeines Verständnis von Datenqualität zu entwickeln, wurde begonnen, eine Ontologie zu entwickeln, die versucht, die unterschiedlichen Sichten in Beziehung zu bringen. Diese Ontologie wurde auf Grund der Integration von WSI in das OpEN.SC-Repository um Aspekte der Bildqualität erweitert (s. http://bioportal.bioontology.org/ontologies/IDQA) (s. Abb. 6).

In der Ontologie werden die Qualitätsparameter Dimensionen zugeordnet und ein Zusammenhang zum Daten-Lebenszyklus sowie zu den Akteuren im jeweiligen Anwendungskontext hergestellt. Insgesamt wurden 18 verschiedene Qualitätsdimensionen aus der Literatur identifiziert [Wand/Wang 1996, Loshin 2001, Pipino et al. 2002, Sebastian-Coleman 2013] und für das OpEN.SC-Repository spezifiziert (s. Abb. 7).

In die Ontologie wurden alle Qualitätsparameter der TMF-Leitlinie in der ersten und dann modifizierten Version aufgenommen. Jeder Parameter wurde auf eine Dimension der Ontologie abgebildet. In der aktuellen Version der TMF-Leitlinie werden alle dort genannten Qualitätsparameter drei Ebenen zugeordnet: Integrität, Organisation und Richtigkeit.

In den Tabellen 11 bis 13 werden aus der Leitlinie die für Forschungsdaten-Repositories interessanten Qualitätsparameter herausgegriffen, in die Bild- und Daten-Qualitätsontologie eingeordnet und deren Bedeutung diskutiert. Die

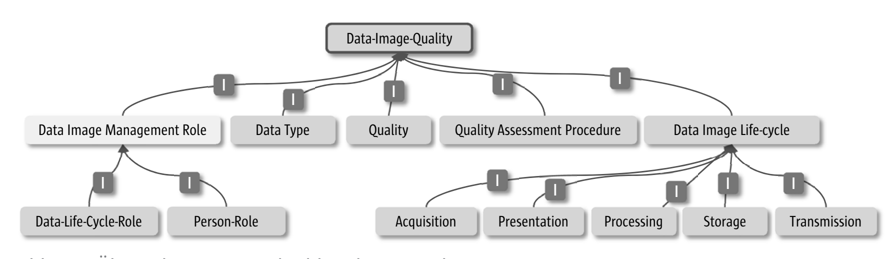



Abb. 6 Übersicht Daten- und Bildqualitätsontologie

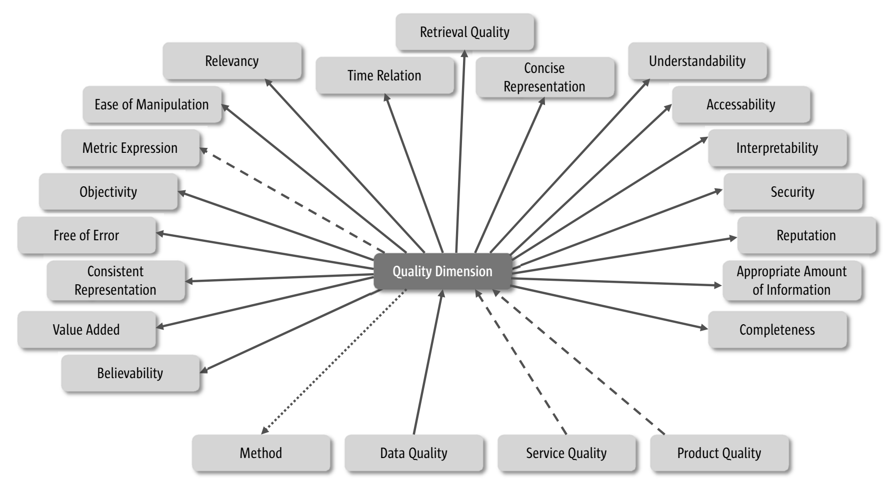



Abb. 7 Dimensionen von Datenqualität in der Ontologie

hier genannten Qualitätsparameter erlauben einen wesentlichen Einblick in die Datenqualität und stellen zwischen den Datenzentren eine Vergleichbarkeit her, die in der gemeinsamen Betrachtungs- und Berechnungsweise begründet ist. Dadurch, dass die möglichen Anfragen nach Daten in Forschungsregistern sehr unterschiedlich sein können, werden weitere Parameter benötigt, die den Datenbestand beschreiben und dabei helfen, dass die Nutzer die Qualität auch bewerten können.

Tab. 11 Qualitätsindikatoren der Ebene Integrität für ein Forschungsdaten-Repository

| TMF-ID | Name des Parameters | Einordnung in die Ontologie |
|---|---|---|
| | Erläuterung | |
| TMF-1003, TMF-1004, TMF-1005 | Widerspruchsfreiheit, sicherer Widerspruch, möglicher Widerspruch | Consistent Representation |
| TMF-1007 | Bevorzugung bestimmter Endziffern | Believability, Interpretability |
| | Das Auftreten von bestimmten Endziffern schränkt sowohl die Glaubhaftigkeit der Daten als auch deren Interpretation ein. In dem Forschungsdaten-Repository kann nicht geklärt werden, warum die Ziffern vergeben wurden. Aus den Häufigkeiten kann lediglich eine Rückkopplung an die Klinik erfolgen, dieses Problem eingehender zu beobachten und ggf. die Datenqualität zu verbessern. | |
| TMF-1035 | Endlosüberlebende | Completeness, Consistent Representation |
| | Die Anzahl der Endlosüberlebenden lässt sich relativ schnell ermitteln. Bei einer hohen Anzahl der Endlosüberlebenden in einem Datensatz ist nicht gleich von einer schlechten Datenqualität auszugehen. Es lässt lediglich die Aussage zu, dass die Sterbedaten in den Datenbestand nicht oder nur wenig nachgetragen werden. Bei Daten aus einer elektronischen Krankenakte ist das eher die Regel. | |




| TMF-ID | Name des Parameters | Einordnung in die Ontologie |
|---|---|---|
| | Erläuterung | |
| TMF-1013 | Anteil fehlender Werte bei Datenelementen | Completeness |
| | Dieser Wert kann für eine Datenquelle erhoben werden, der dann weiter in mandatorisch (TMF-1014) und optional (TMF-1015) differenziert wird. Die Bewertung dieses Parameters muss allerdings differenziert erfolgen. Da es sich um klinische Daten aus dem normalen Behandlungsprozess handelt, bestehen die Daten aus einer Datenquelle auch aus unterschiedlichen Diagnosen und Diagnosegruppen. Jede Erkrankung hat ihr spezifisches Muster an erhobenen Daten. Fehlende Werte in Datenelementen können auch bedeuten, dass diese Werte für Patienten oder Erkrankungen nicht relevant waren und deshalb nicht erhoben wurden, obwohl das Datenelement an das Forschungsregister mit übertragen wurde. Hier muss im Einzelfall entschieden werden, ob der Wert fehlt obwohl dieser prinzipiell vorhanden und/oder für die Erkrankung interessant war. | |
| TMF-1014 | Anzahl fehlender Werte bei mandatorischen Datenelementen | Completeness |
| | Die Menge der mandatorischen Datenelemente ist im OpEN.SC-Register relativ klein und bezieht sich im Wesentlichen auf die Stammdaten wie Alter, Geschlecht und Diagnosen. | |
| TMF-1015 | Anzahl fehlender Werte bei optionalen Datenelementen | Completeness |
| | Siehe auch Kommentar zu TMF-1013. | |
| TMF-1016 | Anteil von Datenelementen mit dem Wert unbekannt o.ä. | Believability, Interpretability |
| | Dieser Parameter ist verwandt mit den Parametern TMF-1007 und TMF-1025. | |
| TMF-1017 | Datenelemente mit bestehenden Einträgen bei allen Beobachtungseinheiten | Completeness |
| | Dieser Parameter ist wichtig für die Beurteilung, in wie weit in allen Datensätzen aus den unterschiedlichen Quellen gesucht werden kann. Im Forschungsregister von OpEN.SC wurden mit den einzelnen Datenquellen separate Vereinbarungen getroffen, welche Daten zum Register übertragen wurden. Für jede Quelle können also unterschiedliche Datenelemente vorliegen. Die Frage ist nun, mit welchem Datenelement erfolgt eine Suche in allen Daten, weil dieses Datenelement in jeder Quelle zumindest geführt wurde. Ob das jeweilige Datenelement auch einen Wert hat, ist damit noch nicht beantwortet und kann nur über die Erhebung von TMF-1013 bis TMF-1015 erfolgen. | |
| TMF-1018 | Ausreißer in stetigen Datenelementen | Believability |
| | Prinzipiell ist eine Analyse von Daten nach Ausreißern sinnvoll. Sie ist allerdings an die Bedingung geknüpft, ob für den betrachteten (Labor-)Wert auch Metainformationen über das Untersuchungsverfahren vorliegen. Nur mittels dieser Metadaten kann ein geeigneter Grenzwert für Ausreißer festgelegt werden. Deshalb ist das Vorliegen von Metadaten zu Laborwerten essentiell (siehe TMF-1050). | |
| TMF-1019 | Werte, die die Messbarkeitsgrenzen von Verfahren unter- oder überschreiten | Believability |
| | Dieser Qualitätsparameter verhält sich analog zu TMF-1018. Um für die Qualitätsanalyse Grenzwerte festzulegen, muss das Messverfahren genau bekannt sein. | |

| TMF-ID | Name des Parameters | Einordnung in die Ontologie |
|---|---|---|
| | Erläuterung | |
| TMF-1020 | Werte aus Standards | Understandability |
| | Die Verwendung von Standardterminologien und Begriffen gemäß internationaler Klassifikationen erhöht das Verstehen z.B. von Diagnosen. Berücksichtigt werden muss hierbei, das für bestimmte medizinische Domänen einige Klassifikationen zu wenig differenziert den Sachverhalt wiedergeben. Zum Beispiel: Diagnostische Begriffe der Pathologie lassen sich sehr schlecht mit den Begriffen der ICD-10 abbilden, die sich eher an klinischen Entitäten orientiert. Umso wichtiger ist es dann, dass die Begriffe aus der Pathologie gemäß einer vereinbarten Terminologie verwendet werden z.B. SNOMED CT. Problematisch wird es insbesondere dann, wenn Spezialbegriffe verwendet werden, die jeweils nur in einer Datenquelle auftreten. | |
| TMF-1021 | Unerlaubte Werte bei qualitativen Datenelementen | Interpretability |
| TMF-1050 | Anteil der von Untersuchungen übermittelten Metadaten | Completeness, Interpretability |
| | Es ist davon auszugehen, dass in einem Forschungsdatenregister Laborwerte aus unterschiedlichen Laboren mit verschiedenen Messverfahren und -bedingungen vorliegen. Diese Laborwerte können nur dann interpretiert und verglichen werden, wenn die Informationen über die Verfahren vorliegen und darüber eine Vergleichbarkeit wenn möglich hergestellt werden kann. Die Bedeutung von Metadaten für Untersuchungsverfahren und -bedingungen kann hier nur unterstrichen werden und gilt auch für alle bildgebenden Verfahren einschließlich der virtuellen Mikroskopie. Eine differenzierte, vergleichende Untersuchung von Laborwerten und Bildern über verschiedene Institutionen hinweg kann nur erfolgen, wenn sichergestellt ist, dass die Werte oder Bilder auch wirklich vergleichbar sind. Deshalb wurde dieser Qualitätsparameter in der Ontologie auch unter Completeness eingeordnet. Für jeden Laborwert bzw. auch für jede Bildkategorie kann genau festgelegt werden, welche Metadaten vorhanden sein müssen, damit die Interpretierbarkeit sichergestellt werden kann. Die Verwendung von LOINC-Kodes ist dabei sehr sinnvoll. | |

Tab. 12 Qualitätsindikatoren der Ebene Organisation für ein Forschungsdaten-Repository

| TMF-ID | Name des Parameters | Einordnung in die Ontologie |
|---|---|---|
| | Erläuterung | |
| TMF-1028 | Aktualität der gespeicherten Daten | Timerelation/Timelines |
| | Im OpEN.SC Forschungsdaten-Repository wird dieser Parameter anders angegeben als in der Leitlinie: Das Forschungsdaten-Repository hat einen regelmäßigen Update-Zyklus (z.B. wöchentlich). Gemäß dieser Frequenz liegen die aktuellen Werte im Register vor. Der Bezug auf einzelne Werte mit Zeitdifferenz und akzeptabler Zeitdifferenz ist so in einem Forschungdatensrepositorium nicht umzusetzen. In der Literatur wird die Verwendung der Begriffe Timelines und Currency diskutiert [Wand/Wang 1996], wobei hier unterschieden wird zwischen der Aktualität der Daten selbst und der Aktualisierung eines Datensatzes. In einem Forschungsregister kann lediglich zur Häufigkeit der Updates aus den Datenquellen Stellung genommen werden. Die Bewertung dieses Parameters ist dann sehr davon abhängig, welche Anforderungen die Nutzerin der Daten hat. | |



| TMF-ID | Name des Parameters | Einordnung in die Ontologie |
|---|---|---|
| | Erläuterung | |
| TMF-1030 | Rekrutierungsrate | Completeness/Population completeness |
| | Dieser Parameter ist der Qualitätsdimension Vollständigkeit zugeordnet, wobei hier der Aspekt der vollständigen Erfassung der gewünschten Personen im Vordergrund steht. Das kann in einem Forschungsdaten-Repository bei der Festlegung der übertragenen Daten und Metadaten bestimmt werden. Später lässt sich dieser Parameter aus dem Datenbestand selbst nicht mehr bestimmen, da keine Relation zur Grundgesamtheit hergestellt werden kann. | |
| TMF-1036 | Anzahl Synonyme | Understandability |
| | Aus Sicht eines Forschungsdaten-Repositories ist es wichtig, die Anzahl möglicher Synonyme bzw. die Synonyme für einen Begriff selbst zu kennen. Eine hohe Anzahl von Synonymen spricht nicht zwangsläufig für eine schlechte Datenqualität. Verwendete, aber nicht bekannte Synonyme erschweren das Verständnis eines Datensatzes. Wenn Synonyme in Klassifikationen und Terminologien benannt sind und entsprechend eingeordnet werden, sind Synonyme unproblematisch. Deshalb ist hier auch der Qualitätsparameter TMF-1020 (Werte aus Standards) interessant. Die Bewertung von Synonymen in einem Datenbestand ist nicht einfach: Das Management von Synonymen ist häufig eher ein technisches Problem und spielt in der Lebenszyklus-Phase der Datenverarbeitung eine Rolle. Die Verarbeitung von Synonymen ist dann kompliziert, wenn sie sich nicht eindeutig auf einen Begriff zuordnen lassen. Somit ist die Anzahl der nicht eindeutig referenzierbaren Synonyme ein interessanter Qualitätsparameter (klassisches Beispiel: HWI als Hinterwandinfarkt und als Harnwegsinfekt), weil dieser auch die Retrievalqualität stark beeinflusst. | |

Tab. 13 Qualitätsindikatoren der Ebene Richtigkeit für ein Forschungsdaten-Repository

| TMF-ID | Name des Parameters | Einordnung in die Ontologie |
|---|---|---|
| | Erläuterung | |
| TMF-1043 | Genauigkeit der Registerdaten | Free of Error |
| | Im OpEN.SC Forschungsdaten-Repository ist diese Frage insbesondere wichtig für Diagnosen. Problematisch dabei aber ist, dass Diagnosen stets Abstraktionsvorgänge sind, die von Ärztin zu Ärztin unterschiedlich sein können. Die Frage nach der wirklichen Diagnose lässt sich nicht beantworten. Hier hat sich das Vorgehen bewährt, dass zu Diagnosen und Fällen insgesamt eine Zweitmeinung eingeholt wird. Damit kann ein Maß angegeben werden, in wie weit Erstmeinung und Zweitmeinung übereinstimmen. Aus diesem Vorgehen heraus lassen sich weitere Maße ableiten: Anzahl der Fälle, für die eine Referenzmeinung eingeholt wurde, Anzahl der übereinstimmenden Fälle, Anzahl der Abweichungen, Anzahl der Fälle mit groben Abweichungen u.a. | |

| TMF-ID | Name des Parameters<br>Erläuterung | Einordnung in die Ontologie |
|---|---|---|
| TMF-1047 | Übereinstimmung mit Verfahrensregeln | Reputation |
| | In einem Forschungsdaten-Repository erfolgt die Auswertung retrograd. Da die Daten aus verschiedenen Quellen kommen und auch der Nutzungskontext offen ist, stellt sich nach der Suche von Fällen die Frage, ob die Qualität der Daten dieser Fälle geeignet ist, um eine Studie durchzuführen (siehe Anwendungsbeispiel). Dieses Maß ist aus Sicht der Forschungsregister ein zusammengesetztes Maß, weil die Verfahrensregeln sehr unterschiedlich sein können. Sie können die Vollständigkeit aller Dateneinträge ebenso fordern (siehe TMF-1013 bis TMF-1015) wie die Kontrolle der Daten und/oder Diagnosen von einer beauftragten Stelle. Die Erfüllung dieser Kriterien ist auf jeden Fall ein gutes Zeichen für Datenqualität und erhöht deren Reputation. | |

## VI.4.4 Beispielhafte Implementierung von Qualitätsindikatoren im OpEN.SC-Forschungsdaten-Repository

### VI.4.4.1 Einführung

Im Rahmen des OpEN.SC-Projektes wurden verschiedene Aspekte der Datenqualität in eigenständigen Forschungsvorhaben untersucht und einige Indikatoren evaluiert. Zunächst wurden die Metadaten von radiologischen Bildern analysiert, die stellvertretend für Labordaten betrachtet werden können (TMF Qualitätsparameter TMF-1050: Anteil der von Untersuchungen übermittelten Metadaten). Für die Entwicklung eines Studyadvisors wurden das Qualitätskriterium TMF-1020: Werte aus Standards eingehender untersucht [Streit et al. 2012; Streit 2013].

### VI.4.4.2 Die Anwendung des Qualitätsfaktors: Umfang der Metadaten bei Untersuchungen

Metadaten enthalten Informationen über die Daten selbst. Bei Laboruntersuchungen sind dem reinen Laborwert selbst weitere Information zugeordnet wie zum Beispiel:

- Welche Labormethode wurde zur Bestimmung des Wertes herangezogen?
- Was sind die Grenzwerte für diese Laboruntersuchung[7]?
- Unter welchen Bedingungen wurde die Untersuchung durchgeführt, z.B. nüchtern oder nach einer Mahlzeit?

7 Bei einigen Laborverfahren werden von Labor zu Labor unterschiedliche Grenzwerte angegeben. Zur Beurteilung eines Wertes ist es allerdings notwendig, die Grenzen zu kennen.

Diese Informationen können mittels eines Kodes der Logical Observation Identifiers Names and Codes (LOINC) als Metadatum übermittelt und gespeichert werden (s. http://loinc.org/). Für radiologische Bilder gibt es ebenfalls Metadaten, die die Untersuchungsverfahren und -bedingungen näher beschreiben, wenn die Bilder im DICOM-Format abgespeichert werden. DICOM kann alle notwendigen Metainformationen zu einem Bild aufnehmen und für die spätere Wiederverwendung bereitstellen (s. http://medical.nema.org/).

In 3.461 anonymisierten DICOM-Bildern aus unterschiedlichen Einrichtungen und von verschiedenen Geräten wurde untersucht, wie viele Metadaten-Items für jedes Bild abgespeichert wurden. Weiterhin wurde dann bestimmt, wie viele Metadaten-Items von allen Bildern verwendet wurden. Die Anzahl gemeinsamer Elemente ist ein wichtiger Indikator dafür, ob eine Suchanfrage auf diesem Bilddatensatz auch wirklich alle Bilder erreicht. In den untersuchten Bildern wurden zwischen 60 und 205 Metadaten-Items gefunden. Am häufigsten wurden die Bilder mit 96 Metadaten versehen (s. Abb. 8).

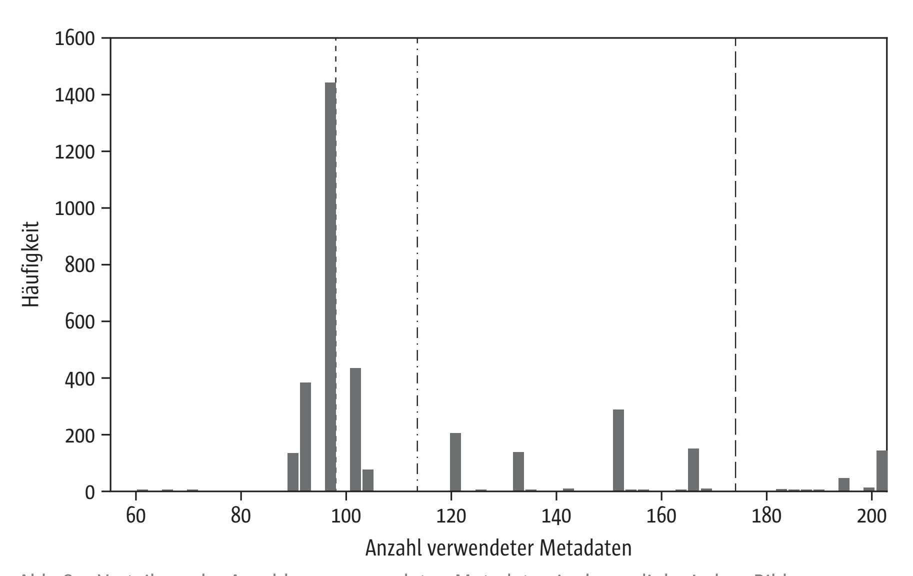



Abb. 8 Verteilung der Anzahl von verwendeten Metadaten in den radiologischen Bildern

Die Anzahl der gemeinsam verwendeten Metadaten war sehr gering: nur 5 Metadaten wurden von allen Bildern verwendet und erlauben damit eine Suche in allen Bildern. Mit nur 15 Metadaten werden 50% der Bilder in einer Suchanfrage erreicht (s. Abb. 9). Mit dieser Untersuchung konnte gezeigt werden, dass selbst die Verwendung eines Standards (hier DICOM) nicht sicherstellt, dass alle Elemente in einem Suchraum abgefragt werden können.

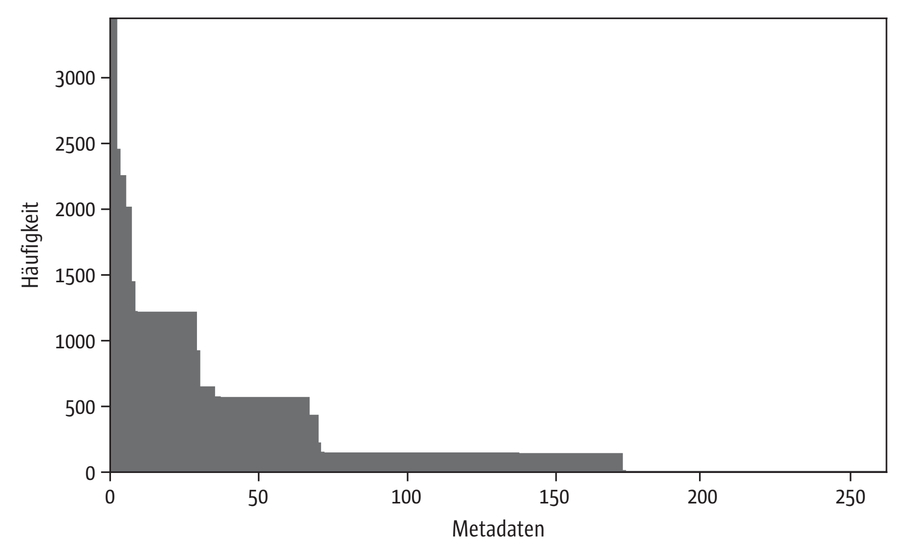



Abb. 9 Häufigkeit gemeinsam verwendeter Metadaten-Items

### VI.4.4.3 Der StudyAdvisor der OpEN.SC-Plattform

Die Durchführung einer empirischen Studie stellt einen komplexen Prozess dar. Das Forschungsdaten-Register des OpEN.SC bietet die Möglichkeit, retrospektiv Untersuchungen an den Daten durchzuführen. Allerdings sind für diese Studien wichtige Kriterien der Datenqualität zu erfüllen.

Der Prozess der Vorbereitung einer retrospektiven Studie schließt die Suche nach geeigneten Patienten und Patientinnen an Hand von Einschluss- und Ausschlusskriterien ein. Am häufigsten wird nach einer bestimmten Diagnose bzw. einer Diagnosekombination, Alter und Geschlecht gesucht. Für das Finden geeigneter Fälle ist es daher sehr wichtig, dass entweder die Diagnosen einen einheitlichen Schlüssel verwenden, z.B. die ICD-10, oder die diagnostischen Begriffe des ICD-10. Werden Synomyma verwendet und lassen sich diese nicht auf die terminologischen Begriffe abbilden, werden diese Fälle nicht gefunden oder/und falsch zugeordnet.

Der StudyAdvisor des OpEN.SC-Projektes ist ein Werkzeug zur Unterstützung der Durchführung von Studien, welches eine Analyse der Datenqualität einschließt (s. Abb. 10). Neben dem Kritierum der Verwendung von Standandbegriffen, die die Interpretierbarkeit von Befunden beeinflussen, wurden Aktualität der Daten und zeitliche Häufigkeit von Untersuchungen (Timedistance aus der Dimension Timerelation) bestimmt. Das Tool integriert sich in die SOA des OpEN.SC-Systems. Über eine Zwischenschicht Hibernate wird auf die OpEN.SC-Datenbank zugegriffen. Die Analyse selbst wird von verschiedenen Modulen ausgeführt (Patientenstammdatenanalyse, Diagnoseanalyse, Laboruntersuchungsanalyse). Die Analyse erfolgt via Zugriff auf die Ontologie zur Bild- und Datenqualität.




Im untersuchten Datenpool von ca. 10.000 Patienten und Patientinnen wurden 90.000 ICD-Kodes in die entsprechenden Felder eingetragen. Demgegenüber stehen aber 110.000 diagnostische Bezeichner. In der weiteren Analyse wurden folgende Probleme gefunden:

- In den ICD-Kodes wurden häufig Xxx.9 („nicht näher bezeichnet") gefunden, was eine ungenaue diagnostische Bezeichnung darstellt und die Interpretierbarkeit deutlich herabsetzt (siehe auch TMF-1007, TMF-1016, TMF-1025). Vor allem aber wird die Glaubwürdigkeit der Daten herabgesetzt, weil dieser Parameter darauf hindeutet, dass die Kodierqualität eingeschränkt ist.
- Beispielsweise wurden bei einem Patienten neben der Hauptdiagnose einer chronischen Pyelonephritis folgende ICD-10 Kodes vergeben:
  - J06.9 – Begriff des Standards: akute Infektion der oberen Atemwege, nicht näher bezeichnet. Bei diesem Patienten wurde stets ein „Infekt, grippal" angegeben.
  - T86.9 – Begriff des Standards: Versagen und Abstoßung/Rejektion eines nicht näher bezeichneten transplantierten Organes oder Gewebes. Hier wurden verschiedene Bezeichner angegeben: Transplantatruputur, Rejektion oder Rejektion (Banff IIb).
  - B18.2 – Begriff des Standards: Chronische Virushepatitis. Hierwurde immer „Hepatitis C, chronisch" als nachvollziehbares Synonym angegeben.

Für die Durchführung einer Studienrekrutierung sind einheitliche Bezeichner bzw. eindeutige Verschlüsselungen notwendig, um sicher potenzielle Kandidaten und Kandidatinnen zu identifizieren. Auch bei der Musteranalyse führt die Verwendung einer Vielzahl von verschiedenen Begriffen dazu, dass keine lokalen Häufigkeiten und Zusammenhänge detektiert werden können.

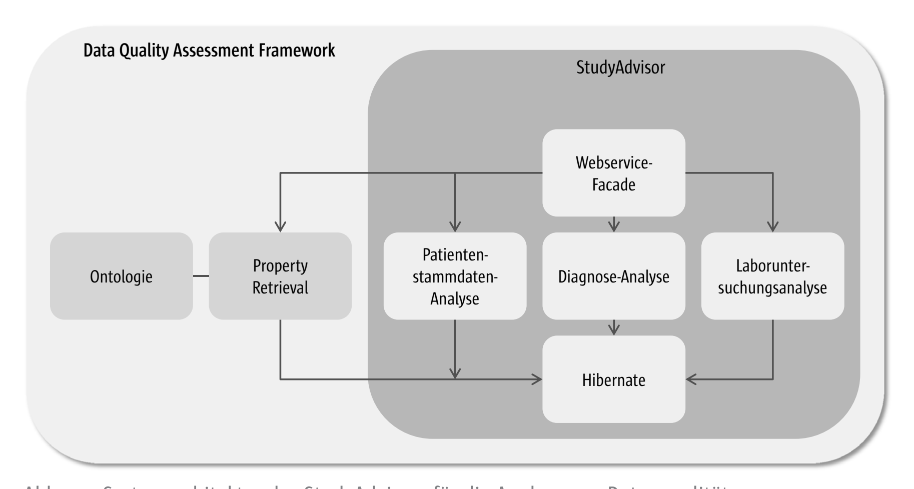



Abb. 10 Systemarchitektur des StudyAdvisors für die Analyse von Datenqualität

## VI.4.5 Zusammenfassung und Ausblick

Für die sekundäre Nutzung von klinischen Daten im Rahmen von Forschungsregistern ist die Darstellung der Datenqualität aus zwei Gründen sehr wichtig:

1. Die Datenqualitätsparameter vermitteln den Nutzern einen Eindruck darüber, mit welchen Daten und in welcher Qualität sie rechnen können. Die Bewertung darüber, ob die Daten nützlich oder für eine bestimmte Studie nicht geeignet sind, hängt sehr stark von der Aufgabenstellung ab und muss den Nutzern überlassen werden.
2. Für eine Reihe von Parametern lässt sich jedoch die Datenqualität unabhängig von deren Nutzung bewerten. So lassen sich daraus Rückschlüsse ziehen, an welchen Stellen Einfluss genommen werden kann. Dabei ist die Unterscheidung von Prozess-, Struktur- und Ergebnisqualität sehr wichtig.

Die Verantwortung für Datenqualität lässt sich nicht immer nur auf den Erzeuger der Daten beschränken. Einige Parameter werden in den anderen Phasen des Datenlebenszyklus beeinflusst. Mit der Leitlinie der TMF ist den Nutzern wie Administratoren ein wichtiges Werkzeug in Hand gegeben worden, um Aussagen über die Datenqualität zu treffen. Für die Forschungsdaten-Repositories und in den aktuellen Entwicklungen von Big Data, wo sehr große Datenmenten (im Terra-, Penta- und Exabyte-Bereich) gesammelt, analysiert und wieder verwendet werden sollen, muss davon ausgegangen werden, dass weitere Qualitätsparameter notwendig sind.

# VII (Implementierungshilfen)

Auf die Bereitstellung von Checklisten wird in Version 2.0 der Leitlinie verzichtet. Die Checklisten aus Version 1.0 stehen im PDF-Format unter www.tmf-ev.de/Leitlinie-Datenqualitaet zum Download zur Verfügung.

# VIII Evaluation

Zu der empfohlenen Vorgehensweise bei der Entwicklung von Leitlinien gehört deren Evaluation [AWMF/ÄZQ 2001]. Hierzu ist eine externe Evaluation der Leitlinie in Form einer prospektiven und kontrollierten Untersuchung erstrebenswert. Zielgröße ist hierbei die Datenqualität, bestimmt einerseits über die einzelnen Qualitätsindikatoren sowie andererseits über den Qualitätsscore. Einzelne Berichte zur Anwendung der Leitlinie in Version 1.0 wurden veröffentlicht [Jacke et al. 2012a, Jacke et al. 2012b, Prokein et al. 2011]. Eine systematische Evaluation der Leitlinie steht noch aus.

# IX Gültigkeitsdauer

Version 2.0 der Leitlinie wurde nach abschließender Diskussion und Konsentierung mit den Projektpartnern am 23. März 2014 fertig gestellt. Neben einer regelmäßigen Aktualisierung sind für die Weiterentwicklung drei Aufgaben relevant. Zum einen steht eine Evaluation im Rahmen einer empirischen Untersuchung noch aus, in deren Anschluss Überarbeitungsbedarf zu erwarten ist. Zum anderen sollte an eine Erweiterung auf andere empirische Forschungsvorhaben wie klinische Studien, weitere Datenbestände der Forschung wie Biomaterialbanken sowie Datenbestände der Versorgung gedacht werden. Letzteres ist dann sowohl für Aufgaben der Versorgung wie Qualitätsmanagement, Projekte der Versorgungsforschung als auch für öffentliche Aufgaben der Surveillance z.B. im Bereich der Arzneimittelsicherheit oder des Infektionsschutzes relevant. Auch das Stichwort „Big Data“ könnte dann aufgenommen werden. Zum dritten deutet sich methodisch eine Ausweitung der Indikatoren zur Qualität von Metadaten an. Hierbei kann an das Nationale Metadata Repository der TMF angedockt werden [Löbe et al. 2012].

*Version:* 2.0
*Erstellungsdatum:* 23.3.2014
*Letzte Überarbeitung:* 23.3.2014
*Überarbeitung geplant:* Spätestens fünf Jahre nach Veröffentlichung.

Verantwortlich für die Fortschreibung der Leitlinie ist die TMF.

# X Autoren

## X.1 Art der Beteiligung

*Literatursichtung, Formulierung der Leitlinie*

Dr. rer. medic. Dipl.-Inf. Michael Nonnemacher, Projektleitung, Institut für Medizinische Informatik, Biometrie und Epidemiologie (IMIBE) am Universitätsklinikum Essen

Dipl.-Bioinf. Daniel Nasseh, Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie (IBE), Ludwig-Maximilians-Universität (LMU) München

Prof. Dr. med. Jürgen Stausberg, Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie (IBE), Ludwig-Maximilians-Universität (LMU) München

*Kooperationspartner*

- Kompetenznetz Angeborene Herzfehler, Berlin:
  - Dr. med. Ulrike Bauer, Geschäftsführerin
- Study of Health in Pomerania (SHIP), Klinisch-Epidemiologische Forschung:
  - Dr. Carsten Oliver Schmidt, Institut für Community Medicine-SHIP/KEF, Greifswald
- Resource Center OpEN.SC:
  - Prof. Dr. Thomas Schrader, Fachhochschule Brandenburg, Fachbereich Informatik & Medien

*Beteiligung am vorangehenden TMF-Projekt V020-04*

Dr. Ron Pritzkuleit, Institut für Krebsepidemiologie e.V., Universität Lübeck

*Entgegennahme der Projektstatusberichte*

Arbeitsgruppe IT-Infrastruktur und Qualitätsmanagement der TMF

## X.2 Finanzielle Unterstützung

Die Erstellung der Leitlinie in Version 2.0 wurde durch die TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. gefördert. Fördermittel erhielten das IMIBE und das IBE für die Literatursichtung und die Überarbeitung der Leitlinie sowie die Projektpartner für die Formulierung der Anwendungsempfehlungen.

## X.3 Interessenkonflikte der Autoren

Es gibt keine Interessenkonflikte der Autoren.

# XI Glossar

| | |
|---|---|
| **Anforderung** | Erfordernis oder Erwartung, das oder die festgelegt, üblicherweise vorausgesetzt oder verpflichtend ist. „Üblicherweise vorausgesetzt" bedeutet, dass es für die Organisation, ihre Kunden und andere interessierte Parteien üblich oder allgemeine Praxis ist, dass das entsprechende Erfordernis oder die entsprechende Erwartung vorausgesetzt ist. Ein Bestimmungswort darf verwendet werden, um eine spezifische Anforderungsart zu bezeichnen, z.B. Qualitätsanforderung. (GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Ausreißer** | Wert eines Datenelements, der nicht den Erwartungen entspricht, d.h. außerhalb des Intervalls der erwarteten oder plausiblen Werte liegt. |
| **Ausschlusskriterien** | Menge von Eigenschaften eines Patienten/einer Person, die seine/ihre Aufnahme in das Register nicht erlauben. Das Vorhandensein von mindestens einer Eigenschaft verbietet die Aufnahme des Patienten/der Person in das Register (z.B. Alter < 18 Jahre). |
| **Data Repository** | Ein Dienst, der Daten aus unterschiedlichen Quellen für Forschungszwecke zugänglich macht. |
| **Datenqualität** | Grad, in dem eine Menge von Daten Anforderungen erfüllt. |
| **Dublette** | Zwei oder mehr Informationseinheiten, die dasselbe Subjekt oder Objekt beschreiben und bei denen die Identität nachgewiesen werden kann oder – sofern der Nachweis nicht möglich ist – mit hinreichender Wahrscheinlichkeit angenommen werden kann. |
| **Einschlusskriterien** | Menge von Eigenschaften eines Patienten/einer Person, die alle gleichzeitig vorhanden sein müssen, damit der er/sie in das Register aufgenommen werden kann (z.B. Vorliegen einer bestimmten Erkrankung). |
| **Endziffer** | Die am weitesten rechts stehende Ziffer eines Wertes. Dabei spielt die Stellung des Kommas keine Rolle (z.B. die Werte „10,43" und „213" besitzen die Endziffer „3"; der Wert „213,0" besitzt die Endziffer „0"). |
| **Ergebnisqualität** | Grad, in dem ein Satz inhärenter Merkmale der Struktur Anforderungen erfüllt. (GMDS-AG Qualitätsmanagement in der Medizin 2003) |
| **Fachliche Qualifikation** | Alle für die Erhebung und Erfassung der Registerdaten notwendige Kenntnisse und Fertigkeiten. Diese umfassen mindestens die detaillierte Kenntnis aller Datenelemente des Registers und der Regeln zu ihrer Abstraktion sowie der für die Datenerfassung eingesetzten EDV-Systeme. |

| | |
|---|---|
| **Genauigkeit** | Das Ausmaß, in dem die Registerdaten der Wahrheit entsprechen. [Arts et al. 2002a] |
| **Hinreichende Wahrscheinlichkeit** | Eine unter Berücksichtigung von Datenqualitätsaspekten und beabsichtigter Datennutzung als ausreichend angesehene Wahrscheinlichkeit. |
| **Homonym** | Mit einem Wort sind zwei unterschiedliche Definitionen verbunden. Im Zusammenhang mit Registern wird ein Patientenidentifikator als Homonym bezeichnet, wenn er mehr als einem Patienten/einer Person zugewiesen wurde. |
| **Kohorte** | Gruppe von Personen, die über eine bestimmte Zeit verfolgt werden, z.B. in einer Kohortenstudie [Porta 2008]. Die Bezeichnungen Kohorte und Kohortenstudie werden in der Leitlinie synonym benutzt. |
| **Konkordanz** | Die Übereinstimmung des Wertes eines Datenelements mit einer Referenzquelle (z.B. Patientenakte, CRF). |
| **Korrelation** | Vorhandensein einer statistischen Beziehung zwischen zwei oder mehr quantitativen Datenelementen. |
| **Metadaten** | Daten über Daten [Berners-Lee 1997]. Im Zusammenhang mit Registern und Kohortenstudien sind hier die Angaben über vorhabensspezifische Charakteristika gemeint. |
| **Originaldaten** | Die Daten, aus denen die Registerdaten extrahiert wurden. Die Originaldaten können in Papierform oder als elektronische Dokumente existieren. |
| **Prozessqualität** | Grad, in dem ein Satz inhärenter Merkmale der Prozesse Anforderungen erfüllt. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualität** | Grad, in dem ein Satz inhärenter Merkmale Anforderungen erfüllt. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitatives Datenelement** | Ein Datenelement, das nur eine endliche Zahl qualitativer Werte annehmen kann, d.h. Werte, die sich nicht quantifizieren lassen (z.B. nach ICD-10 kodierte Diagnose). |
| **Qualitätsbewertung** | Tätigkeit zur Ermittlung der Eignung, Angemessenheit und Wirksamkeit der Betrachtungseinheit, festgelegte Qualitätsziele zu erreichen. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitätsindikator** | Ein Indikator ist ein quantitatives Maß, welches zum Monitoring und zur Bewertung der Qualität wichtiger Leitungs-, Management-, klinischer und unterstützender Funktionen genutzt werden kann, die sich auf das Behandlungsergebnis beim Patienten auswirken. Ein Indikator ist kein direktes Maß der Qualität. Er ist mehr ein Werkzeug, das zur Leistungsbewertung benutzt werden kann, das Aufmerksamkeit auf potenzielle Problembereiche lenken kann, die einer intensiven Überprüfung innerhalb einer Organisation bedürfen könnten. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitätskontrolle** | Siehe Qualitätslenkung. |
| **Qualitätslenkung** | Teil des Qualitätsmanagements, der auf die Erfüllung von Qualitätsanforderungen gerichtet ist. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |



| | |
|---|---|
| **Qualitätsmanagement** | Aufeinander abgestimmte Tätigkeiten zum Leiten und Lenken einer Organisation bezüglich Qualität. Dazu gehören üblicherweise das Festlegen der Qualitätspolitik und der Qualitätsziele, die Qualitätsplanung, die Qualitätslenkung, die Qualitätssicherung und die Qualitätsverbesserung. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitätsmerkmal** | Inhärentes Merkmal eines Produkts, Prozesses oder Systems, das sich auf eine Anforderung bezieht. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitätsplanung** | Teil des Qualitätsmanagements, der auf das Festlegen der Qualitätsziele und der notwendigen Ausführungsprozesse sowie der zugehörigen Ressourcen zur Erfüllung der Qualitätsziele gerichtet ist. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitätssicherung** | Teil des Qualitätsmanagements, der auf das Erzeugen von Vertrauen darauf gerichtet ist, dass Qualitätsanforderungen erfüllt werden. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitätsverbesserung** | Teil des Qualitätsmanagements, der auf die Erhöhung der Fähigkeit zur Erfüllung der Qualitätsanforderungen gerichtet ist. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitätsziele** | Etwas bezüglich Qualität Angestrebtes oder zu Erreichendes. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Qualitätszirkel** | Ein Qualitätszirkel ist eine kleine institutionalisierte Gruppe von fünf bis zwölf Mitarbeitern, die regelmäßig zusammentreffen, um in ihrem Arbeitsbereich auftretende Probleme freiwillig und selbstständig zu bearbeiten. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Register** | Ein Register ist eine möglichst aktive, standardisierte Dokumentation von Beobachtungseinheiten zu vorab festgelegten, aber im Zeitverlauf erweiterbaren Fragestellungen, für die ein präziser Bezug zur Zielpopulation transparent darstellbar ist. [Müller et al. 2010] |
| **Rekrutierung** | Aufnahme von Patienten/Personen, die die Einschluss- und Ausschlusskriterien erfüllen, in ein Register. |
| **Repräsentativität** | Die im Register erfasste Population ist eine Stichprobe der zu untersuchenden Gesamtpopulation, die in wesentlichen Einflussgrößen mit der Gesamtpopulation weitgehend übereinstimmt. |
| **Score** | Messvorschrift für ein Merkmal, die aus einer genau operationalisierten Zusammenfassung von mehreren (i.d.R. mehr als zwei) Komponenten besteht und zu einer mindestens ordinalen Bewertung führt. |
| **Stetiges Datenelement** | Ein Datenelement, das in einem bestimmten Bereich jeden Zwischenwert annehmen kann (z.B. Blutdruck). |
| **Strukturqualität** | Grad, in dem ein Satz inhärenter Merkmale des Ergebnisses Anforderungen erfüllt. [GMDS-AG Qualitätsmanagement in der Medizin 2003] |
| **Studienplan** | Ein oder mehrere Dokumente, die alle wesentlichen Aspekte eines Registers oder einer Kohortenstudie beschreiben (z.B. Einschluss- und Ausschlusskriterien für Fälle, Meldezeiträume, Zeitintervalle für Untersuchungen, Untersuchungs- und Analysemethoden) |

**Synonym** Lexikalisches Zeichen, das die gleiche Bedeutung wie ein anderes lexikalisches Zeichen hat. Im Zusammenhang mit Registern werden Patientenidentifikatoren als Synonyme bezeichnet, wenn sie sich lexikalisch unterscheiden aber demselben Patienten/derselben Person zugewiesen wurden.

**Vollständigkeit** Hier das Ausmaß, in dem die notwendigen Daten, die in das Register eingehen könnten, auch tatsächlich eingegangen sind. [Arts et al. 2002a]

**Vorwert** Ein bezogen auf den aktuellen Wert zu einem früheren Zeitpunkt erhobener Wert desselben Datenelements.

**Vorzeitiges Ausscheiden** Aus der Sicht des Datenmanagements die vorzeitige Beendigung der Datensammlung für einen Patienten/eine Person im Register. Dies kann dadurch geschehen, dass er/sie seine Einwilligung zur Teilnahme am Register zurückzieht oder der Kontakt zwischen Zentrum und Patient/Person abbricht (lost-to-follow-up).

# XII Verzeichnisse

## Literaturverzeichnis


**Abate et al. 1998.** Abate M, Diegert K, Allen H. A hierarchical approach to improving data quality. Data Qual J 1998;4:365–9.

**AGENS 2005.** GPS – Gute Praxis Sekundärdatenanalyse (GPS). 31-1-2005. Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten (AGENS) der Deutschen Gesellschaft für Sozialmedizin und Prävention (DGSMP).

**Akhtar-Danesh/Dehghan-Kooshkghazi 2003.** Akhtar-Danesh N, Dehghan-Kooshkghazi M. How does correlation structure differ between real and fabricated data-sets? BMC.Med Res Methodol. 2003;3:18.

**Allgaier et al. 2005.** Allgaier, S., Ottinger, H., Müller, C. Qualitätsmessung in der medizinischen Dokumentation am Beispiel des Deutschen Registers für Stammzelltransplantationen (DRST). Düsseldorf, Köln: German Medical Science; 2005. http://www.egms.de/en/meetings/gmds2005/05gmds324.shtml [Zugriff am 07.04.2014].

**Altenhofen et al. 2005.** Altenhofen, L., Birkner, B., Blumenstock, G., Geraedts, M., Gibis, B., Jäckel, W., Kopp, I., Kugler, C., Ollenschläger, G., Raspe, H., Reiter, A., Szecsenyi, J., Thoeczek, C., Zorn, U. Qualitätsindikatoren in Deutschland. Positionspapier des Expertenkreises Qualitätsindikatoren beim Ärztlichen Zentrum für Qualität in der Medizin (ÄZQ), Berlin (Stand 14.04.2005). http://www.aezq.de/mdb/edocs/pdf/literatur/qi-positionspapier.pdf [Zugriff am 07.04.2014].

**American Society for Quality o.J.** http://www.asq.org/ [Zugriff am 07.04.2014].

**Arndt et al. 1994.** Arndt S, Tyrrell G, Woolson RF, Flaum M, Andreasen NC. Effects of errors in a multicenter medical study: preventing misinterpreted data. J Psychiatr.Res 1994;28(5):447–59.

**Aronsky/Haug 2000.** Aronsky D, Haug PJ. Assessing the quality of clinical data in a computer-based record for calculating the pneumonia severity index. J.Am.Med.Inform.Assoc. 2000;7(1):55–65.

**Arts et al. 2001.** Arts, D., de Keizer, N., de Jonge, E. Data quality measurement and assurance in medical registries. Patel, V.L. MEDINFO 2001: Proceedings of the 10th World Congress on Medical Informatics. Amsterdam. London: IOS Press; 2001.

**Arts et al. 2002a.** Arts DG, De Keizer NF, Scheffer GJ. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. J.Am.Med.Inform.Assoc. 2002;9(6):600–11.

**Arts et al. 2002b.** Arts D, de Keizer N, Scheffer GJ, de Jonge E. Quality of data collected for severity of illness scores in the Dutch National Intensive Care Evaluation (NICE) registry. Intensive Care Med 2002;28(5):656–9.

**Arts et al. 2003.** Arts DG, Bosman RJ, de Jonge E, Joore JC, De Keizer NF. Training in data definitions improves quality of intensive care data. Crit Care 2003;7(2):179–84.

**AWMF 2000.** Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Erarbeitung von Leitlinien für Diagnostik und Therapie: Methodische Empfehlungen. Februar 2000. http://www.uni-duesseldorf.de/AWMF/ll/ll_metho.htm [Zugriff am 07.04.2014].

**AWMF o.J.** Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Erarbeitung von Leitlinien für Diagnostik und Therapie: Kriterien für die Qualität von Leitlinien. Dez. 2014. http://www.awmf.org/fileadmin/user_upload/Leitlinien/Werkzeuge/Publikationen/methoden.pdf [Zugriff am 21.04.2014].

**AWMF/ÄZQ 2001.** Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Ärztliches Zentrum für Qualität in der Medizin. Das Leitlinien-Manual von AWMF und ÄZQ. Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen 95 (Supplement I) 2001. http://www.awmf.org/leitlinien/awmf-regelwerk/awmf-publikationen-zu-leitlinien/leitlinien-manual.html [Zugriff am 21.04.2014].

**AWMF/ÄZQ 2005.** Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF), Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Fassung 2005/2006. http://www.leitlinien.de/leitlinienmethodik/leitlinienbewertung/delbi [Zugriff am 21.04.2014].

**ÄZQ 2002.** Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). Beurteilung klinischer Messgrößen des Qualitätsmanagements – Qualitätskriterien und -Indikatoren in der Gesundheitsversorgung. Konsenspapier der Bundesärztekammer, der Kassenärztlichen Bundesvereinigung und der AWMF. Z.ärztl.Fortb.Qual.sich. 2002;96(5):2–15.

**Baigent et al. 2008.** Baigent C, Harrell FE, Buyse M, Emberson JR, Altman DG. Ensuring trial validity by data quality assurance and diversification of monitoring methods. Clin Trials. 2008;5(1):49–55.

**Bailey 1991.** Bailey KR. Detecting fabrication of data in a multicenter collaborative animal study. Control Clin Trials 1991;12(6):741–52.

**Bak et al. 2004.** Bak P, Bocker B, Müller W-D, Lohsträter A, Smolenski UC. Zertifizierungs- und Akkreditierungssysteme als Instrument des Qualitätsmanagements in der Rehabilitation – Teil 1 Identifizierung der meist verbreiteten Systeme. Phys Rehab Kur Med 2004;14:243–8.

**Barnett et al. 1978.** Barnett GO, Winickoff R, Dorsey JL, Morgan MM, Lurie RS. Quality assurance through automated monitoring and concurrent feedback using a computer-based medical information system. Med Care 1978;16(11):962–70.

**Barnett/Winickoff 1990.** Barnett GO, Winickoff RN. Quality assurance and computer-based patient records. Am J Public Health 1990;80(5):527–8.

**Barrie/Marsh 1992.** Barrie JL, Marsh DR. Quality of data in the Manchester orthopaedic database. BMJ 1992;304(6820):159–62.

**Baxter et al. 2007.** Baxter C, Dell R, Publ S, Race R. Assessing and improving EHR data quality. J AHIMA 2007;78(3):69–72.

**Benedetto 2003.** Benedetto AR. Six Sigma: not for the faint of heart. Radiol.Manage. 2003;25(2):40–53.

**Benjamin/Watkins 1992.** Benjamin CD, Watkins S. The concern for accurate, quality registry data. J AHIMA 1992;63(3):91–4.

**Berner et al. 2005.** Berner ES, Kasiraman RK, Yu F, Ray MN, Houston TK. Data quality in the outpatient setting: impact on clinical decision support systems. AMIA Annu Symp Proc. 2005:41–5.

**Berners-Lee 1997.** Berners-Lee T. Metadata Architecture. 1997. http://www.w3c.org/DesignIssues/Metadata.html [Zugriff am 07.04.2014].

**Bland/Altmann 1998.** Bland JM, Altman G. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–60

**BMGS 2004.** Bundesministerium für Gesundheit und Soziale Sicherung. Verordnung über die Anwendung der Guten Klinischen Praxis bei der Durchführung von klinischen Prüfungen mit Arzneimitteln zur Anwendung am Menschen (GCP-Verordnung – GCP-V), Bundesgesetzblatt Jahrgang 2004 Teil I Nr. 42, ausgegeben zu Bonn am 12. August 2004. 12-8-2004.

**Bobrowski et al. 1999.** Bobrowski M, Marre M, Yankelevich D. A homogeneous frame-work to measure data quality. In: Proceedings of the 1999 International Conference on Information Quality. Cambridge, MA; 1999.

**Botsis et al. 2010.** Botsis T, Hartvigsen G, Chen F, Weng C. Secondary Use of EHR: Data Quality Issues and Informatics Opportunities. AMIA Summits Transl Sci Proc. 2010;2010:1–5.

**Bray/Parkin 2009.** Bray F, Parkin DM. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. Eur J Cancer. 2009;45(5):747–55.

**Brender et al. 2008.** Brender JD, Suarez L, Langlois PH. Validity of parental work information on the birth certificate. BMC Public Health. 2008;8:95.

**Brennan/Stead 2000.** Brennan PF, Stead WW. Assessing data quality: from concordance, through correctness and completeness, to valid manipulatable representations. J.Am.Med.Inform.Assoc. 2000;7(1):106–7.

**Bronnert et al. 2012.** Bronnert J, Clark JS, Cassidy BS, Fenton S, Hyde L, Kallem C, Watzlaf V. Data quality management model (updated). J AHIMA. 2012;83(7):62–71.

**Brower et al. 2006.** Brouwer HJ, Bindels PJ, Weert HC. Data quality improvement in general practice. Fam Pract. 2006;23(5):529–36.

**Brown et al. 2001.** Brown PJ, Harwood J, Brantigan P. Data quality probes--a synergistic method for quality monitoring of electronic medical record data accuracy and healthcare provision. Medinfo 2001;10(Pt 2):1116–9.

**Brown/Warmington 2002.** Brown PJ, Warmington V. Data quality probes-exploiting and improving the quality of electronic patient record data and patient care. Int J Med Inform 2002;68(1–3):91–8.

**Bundesärztekammer/Kassenärztliche Bundesvereinigung 1997.** Bundesärztekammer, Kassenärztliche Bundesvereinigung. Beurteilungskriterien für Leitlinien in der medizinischen Versorgung. Deutsches Ärzteblatt 1997;33:A 2154-A 2155.

**Bundesärztekammer/Kassenärztliche Bundesvereinigung 2005.** Bundesärztekammer, Kassenärztliche Bundesvereinigung. Das Deutsche Instrument zur methodischen Leitlinienbewertung (DELBI). Deutsches Ärzteblatt 2005;26:B 1616-B 1617.

**Bundeszentrale für gesundheitliche Aufklärung 2001.** Bundeszentrale für gesundheitliche Aufklärung. Qualitätsmanagement in Gesundheitsförderung und Prävention – Grundsätze, Methoden und Anforderungen. Forschung und Praxis der Gesundheitsförderung (15). 2001.

**Califf et al. 1997.** Califf RM, Karnash SL, Woodlief LH. Developing systems for cost-effective auditing of clinical trials. Control Clin.Trials 1997;18(6):651–60.

**Cassidy 1998.** Cassidy B. Vision 2006 brings data quality management into view. J AHIMA. 1998;69(6):28–30.

**Castellion 1983.** Castellion AW. Quality assurance operations: good clinical practices. Drug Inf.J. 1983;17(1):25–9.

**CDC 1999.** Final report: Development, Assessment, and Implementation of an Evaluation for CDC's National Program of Cancer Registries (NPCR). Center for Disease Control and Prevention. 27–7-2005.

**Chiba et al. 2012.** Chiba Y, Oguttu MA, Nakayama T. Quantitative and qualitative verification of data quality in the childbirth registers of two rural district hospitals in Western Kenya. Midwifery. 2012;28(3):329–39.

**Chilvers et al. 1988.** Chilvers CE, Fayers PM, Freedman LS, Greenwood RM, Machin D, Palmer N, Westlake AJ. Improving the quality of data in randomized clinical trials: the COMPACT computer package. COMPACT Steering Committee. Stat.Med. 1988;7(11):1165–70.

**Choquet et al. 2010.** Choquet R, Qouiyd S, Ouagne D, Pasche E, Daniel C, Boussaïd O, Jaulent MC. The Information Quality Triangle: a methodology to assess clinical information quality. Stud Health Technol Inform. 2010;160(Pt 1):699–703.

**Clark et al. 2002.** Clark TG,Altman DG, De Stavola BL. Quantification of the completeness of follow-up. Lancet 2002; 359: 1309–10.

**Couchoud et al. 2013.** Couchoud C, Lassalle M, Cornet R, Jager KJ. Renal replacement therapy registries--time for a structured data quality evaluation programme. Nephrol Dial Transplant. 2013;28(9):2215–20.

**Cnattingius et al. 1990.** Cnattingius S, Ericson A, Gunnarskog J, Kallen B. A quality study of a medical birth registry. Scand.J.Soc.Med. 1990;18(2):143–8.

**Collins 1998.** Collins CM. Ensuring data quality in the pharmaceutical industry. J.AHIMA. 1998;69(6):34–9.

**Collins et al. 1991.** Collins JJ, Buncher CR, Halperin W. Managing the quality and conduct of epidemiologic studies. J.Occup.Med. 1991;33(12):1213–5.

**Cook et al. 2003.** Cook SF, Visscher WA, Hobbs CL, Williams RL. Project IMPACT: results from a pilot validity study of a new observational database. Crit Care Med. 2002;30(12):2765–70.

**Courté-Wienecke 2003.** Courté-Wienecke S. Externe Qualitätssicherung für Managed-Care-Organisationen – Erfahrungen aus den USA. Gesundh ökon Qual manag 2003;8:52–9.

**Crerand et al. 2002.** Crerand W, Lamb J, Rulon V, Karal B, Mardekian J. Building Data Quality into Clinical Trials. JAHIMA. 2002;73:44–56.

**DAE 2004.** Arbeitsgruppe Epidemiologische Methoden der Deutschen Arbeitsgemeinschaft für Epidemiologie (DAE). Leitlinien und Empfehlungen zur Sicherung von Guter Epidemiologischer Praxis (GEP). 2004.

**De 2011.** De S. Hybrid approaches to clinical trial monitoring: Practical alternatives to 100% source data verification. Perspect Clin Res. 2011;2(3):100–4.

**De Lusignan et al. 2002.** De Lusignan S, Stephens PN, Adal N, Majeed A. Does feedback improve the quality of computerized medical records in primary care? J Am Med Inform Assoc 2002;9(4):395–401.

**De Lusignan 2005a.** De Lusignan S. Using feedback to raise the quality of primary care computer data: a literature review. Stud.Health Technol.Inform 2005;116:593–8.

**De Lusignan 2005b.** De Lusignan S. Using the feedback of routinely collected computer data to improve the quality of chronic disease management in UK primary care. Acta Informatica Medica 2005;13(2):59–63.

**DeVor et al. 1992.** DeVor RE, Chang T, Sutherland JW. Statistical quality design and control: contemporary concepts and methods. New York: Macmillan; 1992.

**DIN EN ISO 9001 2000.** DIN EN ISO 9001: Qualitätsmanagementsysteme – Anforderungen. Berlin: Beuth; 2000.

**DIN EN ISO 9004 2000.** DIN EN ISO 9004: Qualitätsmanagementsysteme – Leitfaden zur Leistungsverbesserung. Berlin: Beuth; 2000.

**DIN EN ISO 14050 2010.** DIN EN ISO 14050: Umweltmanagement – Begriffe. Berlin: Beuth; 2010.

**Donabedian 1988.** Donabedian A. The Quality of Care. How Can It Be Assessed? JAMA 1988;260:1743–8.

**Duda et al. 2012.** Duda SN, Shepherd BE, Gadd CS, Masys DR, McGowan CC. Measuring the quality of observational study data in an international HIV research network. PLoS One. 2012;7(4):e33908

**Dyck et al. 2007.** Dyck MJ, Culp K, Cacchione PZ. Data quality strategies in cohort studies: lessons from a study on delirium in nursing home elders. Appl Nurs Res. 2007;20(1):39–43.

**Eberhard/Herrlinger 2010.** Eberhard R, Herrlinger C. Monitoring und Management klinischer Studien. Aulendorf: Editio Cantor Verlag; 2010.

**Elashoff 1999.** Elashoff JD. nQuery Advisor® Version 3.0 User's Guide. Los Angeles, CA; 1999.

**EMEA 2004.** European Medicines Agency (EMEA). CHMP Guideline on data monitoring committees. 2004.

**EMEA 2005a.** European Medicines Agency (EMEA). Procedure for european union guidelines and related documents within the pharmaceutical legislative framework. 2005. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/10/WC500004011.pdf [Zugriff am 21.04.2014].

**EMEA 2005b.** European Medicines Agency (EMEA). ICH Q9 Quality Risk Management. 2005.

**Eminaga et al. 2013.** Eminaga O, Ozgür E, Semjonow A, Herden J, Akbarov I, Tok A, Engelmann U, Wille S. Linkage of Data from Diverse Data Sources (LDS): A Data Combination Model Provides Clinical Data of Corresponding Specimens in Biobanking Information System. J Med Syst 2013;37:9975.

**Europarat 2002.** Europarat. Entwicklung einer Methodik für die Ausarbeitung von Leitlinien für optimale medizinische Praxis. Empfehlung Rec (2001) 13 des Europarates und Erläuterndes Memorandum. 2002. http://www.leitlinien.de/mdb/edocs/pdf/literatur/europaratmethdt.pdf [Zugriff am 21.04.2014].

**Favre 2003.** Favre T. Total Quality Management in Dienstleistungsunternehmen. Konzeptionelle Grundlagen und praktische Umsetzung – dargestellt mit Schwerpunkt Bankwesen [Dissertation]. Universität Basel; 2003.

**Fortier et al. 2011.** Fortier I, Doiron D, Little J, et al. Is rigorous retrospective harmonization possible? Application of the DataSHaPER approach across 53 large studies. Int J Epidemiol 2011;40:1314–28.

**França et al. 2008.** França E, de Abreu DX, Rao C, Lopez AD. Evaluation of cause-of-death statistics for Brazil, 2002–2004. Int J Epidemiol. 2008;37(4):891–901.

**Gassman et al. 1995.** Gassman JJ, Owen WW, Kuntz TE, Martin JP, Amoroso WP. Data quality assurance, monitoring, and reporting. Control Clin.Trials 1995;16(2 Suppl):104S-36S.

**Gaus 2003.** Gaus, W. Dokumentation und Datenverarbeitung bei klinischen Studien. (Bd. 1). 2003. Norderstedt, BoD GmbH. Schriftenreihe Spektrum Medizinische Dokumentation. Gesellschaft zur Förderung Pädagogischer Forschung e.V.

**Gissler et al. 1995.** Gissler M, Teperi J, Hemminki E, Merilainen J. Data quality after restructuring a national medical registry. Scand.J.Soc.Med. 1995;23(1):75–80.

**Gissler et al. 1996.** Gissler M, Ulander VM, Hemminki E, Rasimus A. Declining induced abortion rate in Finland: data quality of the Finnish abortion register. Int.J.Epidemiol. 1996;25(2):376–80.

**Gissler/Shelley 2002.**Gissler M, Shelley J. Quality of data on subsequent events in a routine Medical Birth Register. Med.Inform.Internet.Med. 2002;27(1):33–8.

**Giuffrida et al. 1999.** Giuffrida A, Gravelle H, Roland M. Measuring quality of care with routine data: avoiding confusion between performance indicators and health outcomes. BMJ 1999;319(7202):94–8.

**Gliklich/Dreyer 2010.** Gliklich RE, Dreyer NA, eds. Registries for Evaluating Patient Outcomes: A User's Guide. 2nd ed. (Prepared by Outcome DEcIDE Center [Outcome Sciences, Inc. d/b/a Outcome] under Contract No. HHSA29020050035I TO3.) AHRQ Publication No.10-EHC049. Rockville, MD: Agency for Healthcare Research and Quality. September 2010.

**GMDS-AG Qualitätsmanagement in der Medizin 2003.** GMDS-Arbeitsgruppe „Qualitätsmanagement in der Medizin". Begriffe und Konzepte des Qualitätsmanagements. Informatik, Biometrie und Epidemiologie in Medizin und Biologie 2003;34:1–61.

**Goertzen/Stausberg 2004.** Goertzen R, Stausberg J. Ein Modell zur computerunterstützten Generierung von medizinischen Dokumentationssystemen. Düsseldorf, Köln: German Medical Science; 2004. http://www.egms.de/en/meetings/gmds2004/04gmds305.shtml [Zugriff am 07.04.2014].

**Goldberg et al. 1980.** Goldberg J, Gelfand HM, Levy PS. Registry evaluation methods: a review and case study. Epidemiol Rev. 1980;2:210–20.

**Grimes et al. 2005.** Grimes DA, Hubacher D, Nanda K, Schulz KF, Moher D, Altman DG. The Good Clinical Practice guideline: a bronze standard for clinical research. Lancet 2005;366(9480):172–4.

**Gross/Fessler 2001.** Gross J, Fessler J. Einführung in das Konzept der kontinuierlichen Qualitätsverbesserung. Hessisches Ärzteblatt 2001;4:178–81.

**Håkansson et al. 2001.** Hakansson I, Lundstrom M, Stenevi U, Ehinger B. Data reliability and structure in the Swedish National Cataract Register. Acta Ophthalmol.Scand. 2001;79(5):518–23.

**Haring et al. 2009.** Haring R, Alte D, Volzke H, et al. Extended recruitment efforts minimize attrition but not necessarily bias. J ClinEpidemiol 2009;62:252–60.

**Hartung et al. 1995.** Hartung J, Elpelt B, Klösener K-H. Statistik. Lehr- und Handbuch der angewandten Statistik. 10. Auflage München: R. Oldenbourg Verlag; 1995.

**Hasford/Staib 1994.** Hasford J, Staib AH. Arzneimittelprüfungen und Good Clinical Practice: Planung, Durchführung und Qualitätssicherung. München: MMV Medizin Verlag; 1994.

**Hassey et al. 2001.** Hassey A, Gerrett D, Wilson A. A survey of validity and utility of electronic patient records in a general practice. BMJ 2001;322(7299):1401–5.

**Haug et al. 1991.** Haug PJ, Frederick PR, Tocino I. Quality control in a medical information system. Med Decis. Making 1991;11(4 Suppl):S57-S60.

**Häussler 1998.** Häussler C. Datenqualität. In: Martin W (Herausgeber). Data Warehousing. Bonn: ITP; 1998: 75–89.

**Hedicke et al. 2005.** Hedicke J, Schäfer T, Raspe H, Schübel N, Gadji D, Dalhoff K. Ergebnisvergleich: „Active" vs. „passive reporting" von Fällen mit ambulant erworbener Pneumonie Ergebnisse aus dem Kompetenznetzwerk CAPNETZ. Düsseldorf, Köln: German Medical Science; 2005. http://www.egms.de/en/meetings/gmds2005/05gmds100.shtml [Zugriff am 07.04.2014].

**Helbling/Weinbeck 2004.** Helbling, H., Weinbeck, T. Human-Failure Mode and Effect Analysis. Diplomarbeit. Zürich; 2004.

**Helfert 2001a.** Helfert M. Spezifikation und Messung der Datenqualität in Data-Warehouse-Systemen. Schloss Reisenburg bei Günzburg; 2001.

**Helfert 2001b.** Helfert M. Managing and measuring data quality in data warehousing. In: Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics. Orlando, Fl.; 2001:55–65.

**Heller et al. 2004.** Heller G, Swart E, Mansky T. Ansatz und erste Analysen aus dem Gemeinschaftsprojekt „Qualitätssicherung mit Routinedaten" (QSR). Der Chirurg 2004;43:M191-M199.

**Helou/Ollenschläger 1998.** Helou A, Ollenschlager G. Ziele, Möglichkeiten und Grenzen der Qualitätsbewertung von Leitlinien. Z Arztl.Fortbild.Qualitatssich. 1998;92(5):361–5.

**Hemminki et al. 1992.** Hemminki E, Teperi J, Tuominen K. Need for and influence of feedback from the Finnish birth register to data providers. Qual.Assur.Health Care 1992;4(2):133–9.

**Hentschel/Katalinic 2008.** Hentschel S, Katalinic A, Hrsg. Das Manual der epidemiologischen Krebsregistrierung. Germering/München: Zuckschwerdt; 2008.

**Hogan/Wagner 1997.** Hogan WR, Wagner MM. Accuracy of data in computer-based patient records. J.Am.Med. Inform.Assoc. 1997;4(5):342–55.

**Holle 1995.** Holle R. Methoden zur Konstruktion und Evaluierung klinischer Scores. Habilitationsschrift. Ruprecht-Karls-Universität Heidelberg; 1995.

**Horbar et al. 1995.** Horbar JD, Leahy KA. An assessment of data quality in the Vermont-Oxford Trials Network database. Control Clin Trials 1995;16(1):51–61.

**Hölzer et al. 2001.** Hölzer S, Wächter W, Dudeck J. Auswahl, Probleme und Perspektiven von Qualitätsindikatoren. Z.ärztl.Fortb.Qual.sich. 2001;95:361–6.

**Hruby et al. 2013.** Hruby GW, McKiernan J, Bakken S, Weng C. A centralized research data repository enhances retrospective outcomes research capacity: a case report. J Am Med Inform Assoc 2013;20:563–567.

**Hüsers 2004.** Hüsers, N. Einführung in die (analytische) Qualitätssicherung. 2004. Dresden.

**Iezzoni 1997.** Iezzoni LI. Assessing quality using administrative data. Ann.Intern.Med 1997;127(8 Pt 2):666–74.

**Igl 1999.** Igl G. Rechtlicher Rahmen der Qualitätssicherung in der Medizin nach dem SGB V. Zeitschrift für Gerontologie und Geriatrie 1999;32:3–10.

**Jacke et al. 2010.** Jacke C, Kalder M, Etter J, Koller M, Mueller UO, Wagner U, Albert US. Optimierung der Datenqualität von Kohortenstudien mittels Qualitätsindikatoren. Monitor Versorgungsforschung 2010;3 (Sonderausgabe).

**Jacke et al. 2012a.** Jacke CO, Kalder M, Koller M, Wagner U, Albert US. Systematische Bewertung und Steigerung der Qualität medizinischer Daten. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2012;55(11–12):1495–503.

**Jacke et al. 2012b.** Jacke CO, Kalder M, Wagner U, Albert US. Valid comparisons and decisions based on clinical registers and population based cohort studies: assessing the accuracy, completeness and epidemiological relevance of a breast cancer query database. BMC Res Notes. 2012;5:700.

**Jarke/Vassiliou 1997.** Jarke M, Vassiliou Y. Data warehouse quality: a review of the DWQ project. 1997; Cambridge MA: 1997.

**JCAHO 1990.** Joint Commission on Accreditation of Healthcare Organizations. Primer on Indicator Development and Application. Measuring Quality in Health Care. Illinois; 1990.

**Jensen et al. 2002.** Jensen AR, Overgaard J, Storm HH. Validity of breast cancer in the Danish Cancer Registry. A study based on clinical records from one county in Denmark. Eur.J.Cancer Prev. 2002;11(4):359–64.

**Jick et al. 1991.** Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. BMJ 1991;302(6779):766–8.

**Jung/Winter 2000.** Jung R, Winter R. Data Warehousing Strategie. Erfahrungen, Methoden, Visionen. Berlin, Heidelberg: Springer; 2000.

**Kahn et al. 2012.** Kahn MG, Raebel MA, Glanz JM, Riedlinger K, Steiner JF. A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. Med Care. 2012;50 Suppl:S21–9.

**Kaomea 1994.** Kaomea, P. Valuation of data quality: a decision analysis approach. 1994. Cambridge, MA, Massachusetts Institute of Technology (MIT) Sloan School of Management. http://web.mit.edu/tdqm/www/papers/94/94-09.html [Zugriff am 07.04.2014]

**Katalinic 2005.** Katalinic A. Qualitätssicherung von Daten in epidemiologischen Krebsregistern. Vortrag gehalten auf dem Workshop Medizinische Daten in Forschung und Versorgung, Essen 16.02.2005 (nicht publiziert).

**Khosla et al. 2000.** Khosla R, Verma DD, Kapur A, Khosla S. Efficient source data verification. Indian J. Pharmacol. 2000; 32(3):180–6.

**Knatterud 1981.** Knatterud GL. Methods of quality control and of continuous audit procedures for controlled clinical trials. Control Clin.Trials 1981;1(4):327–32.

**Knatterud et al. 1998.** Knatterud GL, Rockhold FW, George SL, Barton FB, Davis CE, Fairweather WR, Honohan T, Mowery R, O'Neill R. Guidelines for quality assurance in multicenter trials: a position paper. Control Clin. Trials 1998;19(5):477–93.

**Knaup et al. 2004.** Knaup P, Garde S, Haux R. Ein Meta-Modell für kooperative Dokumentationsumgebungen. Düsseldorf, Köln; 2004. http://www.egms.de/en/meetings/gmds2004/04gmds033.shtml [Zugriff am 07.04.2014].

**Krzych et al. 2011.** Krzych LJ, Lees B, Nugara F, Banya W, Bochenek A, Cook J, Taggart D, Flather MD. Assessment of data quality in an international multi-centre randomised trial of coronary artery surgery. Trials. 2011;12:212.

**Kühn-Stoffers 2004.** Kühn-Stoffers, P. Script zur Vorlesung Grundlagen der Analytik II. Teil 3: Qualitätssicherung in der Analytik. 2004.

**Kuntoro et al. 1994.** Kuntoro, LaPorte RE, Mazumdar S. Approaches to quality control with an application to a new cancer registry in a developing country. J.Clin.Epidemiol. 1994;47(7):779–86.

**Kuperman et al. 1998.** Kuperman GJ, Sussman A, Schneider LI, Fiskio JM, Bates DW. Towards improving the accuracy of the clinical database: allowing outpatients to review their computerized data. Proc.AMIA. Symp. 1998:220–4.

**Lachin 2004.** Lachin JM. Conflicts of interest in data monitoring of industry versus publicly financed clinical trials. Stat.Med 2004;23(10):1519–21.

**Lack/Schneider 2005.** Lack N, Schneider KTM. Qualitätsmessung zwischen Illusion und Realität – Ist Qualität in der Medizin messbar? Z Geburtshilfe Neonatol 2005;209:1–7.

**Lai et al. 2012.** Lai JK, Lucas RM, Banks E, Ponsonby AL, Ausimmune Investigator G. Variability in vitamin D assays impairs clinical assessment of vitamin D status. Intern Med J 2012;42:43–50.

**Landis/Koch 1977.** Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–174.

**Larsen et al. 2009.** Larsen IK, Småstuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, Møller B. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. Eur J Cancer. 2009;45(7):1218–31.

**Leiner/Haux 1996.** Leiner F, Haux R. Systematic planning of clinical documentation. Methods Inf.Med. 1996;35(1):25–34.

**Lin/Chen 2000.** Lin B, Chen H. Improving data quality in the healthcare industry: a framework. In: Proceedings of the 31st Annual Meeting of the Decision Sciences Institute (DSI). Orlando, Fl.; 2000:702–4.

**Lin et al. 2004.** Lin CM, Lee PC, Teng SW, Lu TH, Mao IF, Li CY. Validation of the Taiwan Birth Registry using obstetric records. J Formos.Med Assoc 2004;103(4):297–301.

**Lindquist 2004.** Lindquist M. Data quality management in pharmacovigilance. Drug Saf 2004;27(12):857–70.

**Loane et al. 2011.** Loane M, Dolk H, Garne E, Greenlees R; EUROCAT Working Group. Paper 3: EUROCAT data quality indicators for population-based registries of congenital anomalies. Birth Defects Res A Clin Mol Teratol. 2011;91 Suppl 1:S23–30.

**Löbe et al. 2012.** Löbe M, Drepper J, Ngouongo SMN, Schollmeyer M, Stausberg J, Verplancke P, Löffler M. Das nationale Metadata Repository für die klinische und epidemiologische Forschung (MDR). In: Stausberg J, Großer A, Haerting J, Knaup P, Plischke M, Timmer A, Haux R, Hrsg. GMDS 2012. 57. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V. (GMDS). Abstractband. Stuttgart: Schattauer, 2012: 76–77.

**Logan et al. 2001.** Logan JR, Gorman PN, Middleton B. Measuring the quality of medical records: a method for comparing completeness and correctness of clinical encounter data. Proc.AMIA.Symp. 2001:408–12.

**Lorenzoni et al. 1999.** Lorenzoni L, Da Cas R, Aparo UL. The quality of abstracting medical information from the medical record: the impact of training programmes. Int.J.Qual Health Care 1999;11(3):209–13.

**Loshin 2001.** Loshin D. Enterprise knowledge management: The data quality approach. San Diego: Morgan Kaufmann, 2001. xviii, 493.

**Lu et al. 1995.** Lu JH, Hung JH, Lin FM, Shen WY, Chen SJ, Hwang B, Wu SI, Yu Chao YM. A quality study of a computerized medical birth registry. J Obstet.Gynaecol. 1995;21(1):23–30.

**Lucas 2002.** Lucas J. The Essential Six Sigma. Quality Progress 2002:27–31.

**Lüdemann et al. 2000.** Lüdemann J, Piek M, Wood WG, et al. Methoden zur Qualitätssicherung im medizinischen Untersuchungsbereich epidemiologischer Feldstudien: Die „Study of Health in Pomerania“ (SHIP). Gesundheitswesen 2000;62:234–43.

**Lyberg et al. 1997.** Lyberg L, Biemer P, Collins M, de Leeuw E, Dippo C, Schwarz N, Trewin, D. (Herausgeber). Survey Measurement and Process Quality. New York, Chichester, Weinheim, Brisbane, Singapore, Toronto, John Wiley & Sons. Wiley Series in Probalility ans Statistics; 1997.

**Macaulay et al. 1996.** Macaulay EM, Cooper GG, Engeset J, Naylor AR. Prospective audit of discharge summary errors. Br J Surg 1996;83(6):788–90.

**Macefield et al. 2013.** Macefield RC, Beswick AD, Blazeby JM, Lane JA. A systematic review of on-site monitoring methods for health-care randomised controlled trials. Clin Trials. 2013;10(1):104–24.

**Mainz 2003a.** Mainz J. Defining and classifying clinical indicators for quality improvement. Int J Qual.Health Care 2003;15(6):523–30.

**Mainz 2003b.** Mainz J. Developing evidence-based clinical indicators: a state of the art methods primer. Int J Qual.Health Care 2003;15 Suppl 1:i5–11.

**Malin et al. 2002.** Malin JL, Kahn KL, Adams J, Kwan L, Laouri M, Ganz PA. Validity of cancer registry data for measuring the quality of breast cancer care. J Natl.Cancer Inst. 2002;94(11):835–44.

**Malin/Keating 2005.** Malin JL, Keating NL. The cost-quality trade-off: need for data quality standards for studies that impact clinical practice and health policy. Journal of Clinical Oncology 2005;23(21):4581–4.

**Marco 2000.** Marco D. Building and managing the meta data repository: A full lifecycle guide. New York: Wiley, 2000. xxi, 392.

**Marks 2004.** Marks RG. Validating electronic source data in clinical trials. Control Clin.Trials 2004;25(5):437–46.

**Maruszewksi et al. 2005.** Maruszewski B, Lacour-Gayet F, Monro JL, Keogh BE, Tobota Z, Kansy A. An attempt at data verification in the EACTS Congenital Database. Eur J Cardiothorac.Surg 2005;28(3):400–4.

**Massey/Hoffman 1989.** Massey JT, Hoffman KL. Monitoring data quality through comparisons between data systems. Stat.Med. 1989;8(3):367–77.

**Mathieu/Khalil 1998.** Mathieu RG, Khalil O. Data quality in the database systems course. Data Qual J 1998;4(1).

**McDonald 1999.** McDonald CJ. Quality measures and electronic medical systems. JAMA 1999;282(12):1181–2.

**McKenzie et al. 2005.** McKenzie K, Walker S, Besenyei A, Aitken LM, Allison B. Assessing the concordance of trauma registry data and hospital records. HIM J. 2005;34(1):3–7.

**Messenger et al. 2012.** Messenger JC, Ho KK, Young CH, Slattery LE, Draoui JC, Curtis JP, Dehmer GJ, Grover FL, Mirro MJ, Reynolds MR, Rokos IC, Spertus JA, Wang TY, Winston SA, Rumsfeld JS, Masoudi FA; NCDR Science and Quality Oversight Committee Data Quality Workgroup. The National Cardiovascular Data Registry (NCDR) Data Quality Brief: the NCDR Data Quality Program in 2012. J Am Coll Cardiol. 2012;60(16):1484–8.

**Mettin 2002.** Mettin, P. Qualitäts- und Risikomanagement in Bildungsnetzwerken. Informationsmaterial zum Modul Nr. 4: Bildungsnetzwerke bewerten. 02.09.2002. ANUBA.

**Mikkelsen/Aasly 2001.** Mikkelsen G, Aasly J. Concordance of information in parallel electronic and paper based patient records. Int.J.Med.Inform. 2001;63(3):123–31.

**Moczygemba/Hewitt 2001.** Moczygemba J, Hewitt B. Managing clinical data in an electronic environment. Health Care Manag. (Frederick.) 2001;19(4):33–8.

**Mohr et al. 2005.** Mohr VD, Bauer J, Döbler K, Eckert O, Fischer B, Woldenga C, Hrsg. Qualität sichtbar machen. BQS-Qualitätsreport 2004. Düsseldorf: BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH; 2005: 60–1.

**Montgomery 1991.** Montgomery DC. Introduction to statistical quality control. New York: Wiley; 1991.

**Morganstein/Marker 1997.** Morganstein D, Marker DA. Continuous quality improvement in statistical agencies. In: Lyberg L, Biemer P, Collins M, de Leeuw E, Dippo C, Schwarz N et al., editors. Survey Measurement and Process Quality. New York, Chichester, Weinheim, Brisbane, Singapore, Toronto: John Wiley & Sons, Inc.; 1997: 475–500.

**Mphatswe et al. 2012.** Mphatswe W, Mate KS, Bennett B, Ngidi H, Reddy J, Barker PM, Rollins N. Improving public health information: a data quality intervention in KwaZulu-Natal, South Africa. Bull World Health Organ. 2012;90(3):176–82.

**Müller et al. 2010.** Müller D, Augustin M, Banik N, Baumann W, Bestehorn K, Kieschke J, Lefering R, Maier B, Mathis S, Rustenbach SJ, Sauerland S, Semler SC, Stausberg J, Sturm H, Unger C, Neugebauer EA. Memorandum Register für die Versorgungsforschung. Gesundheitswesen. 2010 Nov;72(11):824–39.

**Mullooly 1990.** Mullooly JP. The effects of data entry error: an analysis of partial verification. Comput.Biomed. Res. 1990;23(3):259–67.

**Nahm et al. 2008.** Nahm ML, Pieper CF, Cunningham MM. Quantifying data quality for clinical trials using electronic data capture. PLoS One. 2008;3(8):e3049.

**Naumann/Rolker 2000.** Naumann F, Rolker C. Assessment methods for information quality criteria. In: Proceedings of the 2000 Conference on Information Quality. Cambridge, MA; 2000: 148–62.

**Neaton et al. 1990.** Neaton JD, Duchene AG, Svendsen KH, Wentworth D. An examination of the efficiency of some quality assurance methods commonly employed in clinical trials. Stat Med 1990;9(1–2):115–23.

**Ngouongo et al. 2013.** Ngouongo, Sylvie M N, Löbe M, Stausberg J. The ISO/IEC 11179 norm for metadata registries: does it cover healthcare standards in empirical research? J Biomed Inform 2013; 46:318–327.




**Nickelsen et al. 2004.** Nickelsen TN, Harling H, Kronborg O, Bulow S, Jorgensen T. The completeness and quality of the Danish Colorectal Cancer clinical database on colorectal cancer. Ugeskr.Laeger 2004;166(36):3092–5.

**Nielsen et al. 1996.** Nielsen GL, Sorensen HT, Pedersen AB, Sabroe S. Analyses of data quality in registries concerning diabetes mellitus–a comparison between a population based hospital discharge and an insulin prescription registry. J.Med.Syst. 1996;20(1):1–10.

**Nielsen et al. 2004.** Nielsen DM, Merry MD, Schyve PM, Bisognano M. Can the gurus' concepts cure healthcare? Quality Progress 2004:25–34.

**Nonnemacher et al. 2007.** Nonnemacher M, Weiland D, Stausberg J. Leitlinie zum adaptiven Management von Datenqualität in Kohortenstudien und Registern Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2007.

**Norton et al. 1981.** Norton SL, Buchanan AV, Rossmann DL, Chakraborty R, Weiss KM. Data entry errors in an on-line operation. Comput.Biomed.Res. 1981;14(2):179–98.

**O'Kelly 2004.** O'Kelly M. Using statistical techniques to detect fraud: a test case. Pharmaceutical Statistics 2004;3:237–46.

**O'Sullivan et al. 1996.** O'Sullivan DP, Needham CA, Bangs A, Atkin K, Kendall FD. Postmarketing surveillance of oral terbinafine in the UK: report of a large cohort study. Br J Clin Pharmacol. 1996;42(5):559–65.

**Ose et al. 2004.** Ose C, Weiland D, Jöckel K-H. Low budget GCP – am Beispiel Monitoring. How to practice low cost monitoring in clinical trials. Informatik, Biometrie und Epidemiologie in Medizin und Biologie 2004;35:54–62.

**Parkin/Muir 1992.** Cancer incidence in five continents. Volume VI. Parkin, D.M., Muir, C.S., Whelan, S.L., Gao, Y.T., Ferlay, J., and Powell, J. (120). Lyon: International Agency for Research on Cancer. IARC Scientific Publications; 1992.

**Pflüglmayer 2001.** Pflüglmayer M. Informations- und Kommunikationstechnologien zur Qualitätsverbesserung im Krankenhaus [Dissertation]. Institut für Wirtschaftsinformatik, Sozial- und Wirtschaftswissenschaftliche Fakultät der Johannes Kepler Universität Linz, Österreich; 2001.

**Pipino et al. 2002.** Pipino LL, Lee YW, Wang RY. Data quality assessment. Commun. ACM 2002;45:211.

**Piwowar/Vision 2013.** Piwowar HA, Vision TJ. Data reuse and the open data citation advantage. PeerJ 20013;1:e175.

**Pogash et al. 2001.** Pogash RM, Boehmer SJ, Forand PE, Dyer AM, Kunselman SJ. Data management procedures in the Asthma Clinical Research Network. Control Clin.Trials 2001;22(6 Suppl):168S-80S.

**Pollock 1994.** Pollock BH. Quality assurance for interventions in clinical trials. Multicenter data monitoring, data management, and analysis. Cancer 1994;74(9 Suppl):2647–52.

**Porcheret et al. 2004.** Porcheret M, Hughes R, Evans D, Jordan K, Whitehurst T, Ogden H, Croft P. Data quality of general practice electronic health records: the impact of a program of assessments, feedback, and training. J Am Med Inform Assoc 2004;11(1):78–86.

**Porta 2008.** Porta M, ed. A Dictionary of Epidemiology. New York: Oxford University Press; 2008.

**Poy 1993.** Poy E. Objectives of QC systems and QA function in clinical research. Qual Assur. 1993;2(4):326–31.

**Pringle et al. 1995.** Pringle M, Ward P, Chilvers C. Assessment of the completeness and accuracy of computer medical records in four practices committed to recording data on computer. Br.J.Gen.Pract. 1995;45(399):537–41.

**Prins et al. 2000.** Prins H, Kruisinga FH, Buller HA, Zwetsloot-Schonk JH. Availability and accuracy of electronic patient data for medical practice assessment. Stud.Health Technol.Inform. 2000;77:484–8.

**Prokein et al. 2011.** Prokein J, Steinmann S, Wiese B, Schäfer I. Bestimmung der Datenqualität und Qualitätsberichterstattung in epidemiologischen Studien am Beispiel der Studie MultiCare 1. In: Mainz//2011. 56. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (gmds), 6. Jahrestagung der Deutschen Gesellschaft für Epidemiologie (DGEpi). Mainz, 26.-29.09.2011. Düsseldorf: German Medical Science GMS Publishing House; 2011. Doc11gmds129. DOI: 10.3205/11gmds129, URN: urn:nbn:de:0183-11gmds1298.

**Rajaraman/Samet 2005.** Rajaraman P, Samet JM. Quality Control and Good Epidemiological Practice. In: Ahrens W, Pigeot I, eds. Handbook of Epidemiology. Bremen: Springer; 2005:503–56.

**Rao et al. 2005.** Rao C, Lopez AD, Yang G, Begg S, Ma J. Evaluating national cause-of-death statistics: principles and application to the case of China. Bull World Health Organ 2005;83:617–25.

**Renard/van Glabbeke 1989.** Renard J, van Glabbeke M. Quality control and data handling. In: Rotmensz N, Vantongelen K, Renard J, editors. Data management and clinical trials. New York: Elsevier; 1989: 147–62.

**Ricketts et al. 1993.** Ricketts D, Newey M, Patterson M, Hitchin D, Fowler S. Markers of data quality in computer audit: the Manchester Orthopaedic Database. Ann.R.Coll.Surg.Engl. 1993;75(6):393–6.

**Roos et al. 1989.** Roos LL, Sharp SM, Wajda A. Assessing data quality: a computerized approach. Soc.Sci.Med 1989;28(2):175–82.

**Roos et al. 2003.** Roos DE, Davis SR, Turner SL, O'Brien PC, Spry NA, Burmeister BH, Hoskin PJ, Ball DL. Quality assurance experience with the randomized neuropathic bone pain trial (Trans-Tasman Radiation Oncology Group, 96.05). Radiother.Oncol. 2003;67(2):207–12.

**Rossi 2005.** Rossi M. Data Management und Datendeskription mit SAS [Diplomarbeit]. Fachhochschule Hannover; 2005.

**Rouff/Child 2003.** Rouff J, Child C. Application of quality improvement theory and process in a national multicenter HIV/AIDS clinical trials network. Qual.Manag.Health Care 2003;12(2):89–96.

**Rubin et al. 2001.** Rubin HR, Pronovost P, Diette GB. From a process of care to a measure: the development and testing of a quality indicator. Int J Qual.Health Care 2001;13(6):489–96.

**Sáez et al. 2012.** Sáez C, Martínez-Miranda J, Robles M, García-Gómez JM. Organizing data quality assessment of shifting biomedical data. Stud Health Technol Inform. 2012;180:721–5.

**Salati et al. 2011.** Salati M, Brunelli A, Dahan M, Rocco G, Van Raemdonck DE, Varela G; European Society of Thoracic Surgeons Database Committee. Task-independent metrics to assess the data quality of medical registries using the European Society of Thoracic Surgeons (ESTS) Database. Eur J Cardiothorac Surg. 2011;40(1):91–8.

**Savage 1999.** Savage AM. Framework for characterizing data and identifying anomalies in health care databases. Proc.AMIA.Symp. 1999:374–8.

**Schmidt et al. 1995.** Schmidt J, Gertzen H, Aschenbrenner KM, Ryholt-Jensen S. Detecting fraud using auditing and biometrical methods. Appl Clin Trials 1995;4(5):40–9.

**Schneider et al. 2003.** Schneider A, Broge B, Szecsenyi J. Müssen wir messen, um (noch) besser werden zu können? Die Bedeutung von Qualitätsindikatoren in strukturierten Behandlungsprogrammen und Qualitätsmanagement. Z Allg Med 2003;79:547–52.

**Schulman/Hahn 1993.** Schulman J, Hahn JA. Quality control of birth defect registry data: a case study. Public Health Rep. 1993;108(1):91–8.

**Sebastian-Coleman 2013.** Sebastian-Coleman L. Measuring data quality for ongoing improvement: A data quality assessment framework. Burlington: Elsevier Science, 2013. Online resource.

**Selbmann 2000.** Selbmann HK. Die sozialmedizinischen Komponenten des Qualitätsmanagements und was die Sozialmedizin aus dem Qualitätsmanagement lernen kann. Vortrag anlässlich der Verleihung der Salomon-Neumann-Medaille der Deutschen Gesellschaft für Sozialmedizin und Prävention 1999. Gesundheitswesen 2000;62:123–6.

**Sens et al. 2007.** Sens B, Fischer B, Bastek A, Eckardt J, Kaczmarek D, Paschen U, Pietsch B, Rath S, Ruprecht T, Thomeczek C, Veit C, Wenzlaff P. Begriffe und Konzepte des Qualitätsmanagements – 3. Auflage. GMS Med Inform Biom Epidemiol. 2007;3(1):Doc05.

**Seufert et al. 2000.** Seufert R, Woernle F, Brockenhoff P, Knapstein PG. Die Integration qualitätssichernder Prozeduren in medizinische Dokumentationssysteme – vom Qualitätsindikator zur verbesserten medizinischen Qualität. Zentralbl Gynakol 2000;122:602–6.

**Sigurdardottir et al. 2012.** Sigurdardottir LG, Jonasson JG, Stefansdottir S, Jonsdottir A, Olafsdottir GH, Olafsdottir EJ, Tryggvadottir L. Data quality at the Icelandic Cancer Registry: comparability, validity, timeliness and completeness. Acta Oncol. 2012;51(7):880–9.

**Siegel/Catellan 1988.** Siegel S, Castellan JN. Nonparametric Statistics for the Behavioral Sciences. 2nd edition McGraw-Hill; 1988.

**Sinaci et al. 2013.** Sinaci AA, Laleci Erturkmen, Gokce B. A federated semantic metadata registry framework for enabling interoperability across clinical research and care domains. J Biomed Inform 2013;46:784–794.

**Shabestari/Roudsari 2013.** Shabestari O, Roudsari A. Challenges in data quality assurance for electronic health records. Stud Health Technol Inform. 2013;183:37–41.

**Slaughter/Willner 1996.** Slaughter J, Willner S. A successful methodology for improving the quality of clinical data. JAHIMA. 1996;67(10):46–1.

**Solomon et al. 1991.** Solomon DJ, Henry RC, Hogan JG, Van Amburg GH, Taylor J. Evaluation and implementation of public health registries. Public Health Rep. 1991;106(2):142–50.

**Sorensen et al. 1996.** Sorensen HT, Sabroe S, Olsen J. A framework for evaluation of secondary data sources for epidemiological research. Int J Epidemiol 1996;25(2):435–42.

**Spyra et al. 2002.** Spyra K, Müller-Fahrnow W, Held K, Karoff M, Nowossadeck E. Indikatoren der Prozess- und Ergebnisqualität im Rahmen eines multizentrischen Qualitätsmanagements in der kardiologischen Rehabilitation. Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen 2002;96:31–6.

**Stausberg 1997.** Stausberg J. Evaluation of quality control methods for medical documentation. Stud.Health Technol.Inform. 1997;43 Pt B:864–8.

**Stausberg/Albrecht 1996.** Stausberg J, Albrecht KH. Unterstützung der standardisierten Basisdokumentation nach dem Gesundheitsstrukturgesetz durch Qualitätszirkel. Abstracts. 41. Jahrestagung der GMDS. Bonn; 1996: 97.

**Stausberg/Albrecht 1997.** Stausberg J, Albrecht KH. Bewertung von Verfahren zur Qualitätssicherung der Basisdokumentation nach dem Gesundheitsstrukturgesetz. Medizinische Informatik, Biometrie und Epidemiologie in Medizin und Biologie 1997;82:137–41.

**Stausberg/Jöckel 2001.** Stausberg J, Jöckel K-H. Integriertes Qualitätsmanagement im Krankenhaus am Beispiel eines Krankenhausinformationssystems für ein operatives Zentrum. In: Schneider A, Pflaumer P, editors. Power Tools. Management-, Beratungs- und Controllinginstrumente. Wiesbaden: Betriebswirtschaftlicher Verlag Dr. Th. Gabler GmbH; 2001: 361–9.

**Stausberg et al. 2003a.** Stausberg J, Bilir H, Waydhas C, Ruchholtz S. Guideline validation in multiple trauma care through business process modeling. Int J Med Inform 2003;70(2–3):301–7.

**Stausberg et al. 2003b.** Stausberg J, Koch D, Ingenerf J, Betzler, M. Qualität der Diagnosen- und Prozedurendokumentation in Krankenakte und rechnergestützter Basisdokumentation. 2. Norddeutsches Symposium für Medizinische Dokumentation und Pflegeberufe. Tagungsband. Hamburg; 2003: 243–244.

**Stausberg et al. 2003c.** Stausberg J, Koch D, Ingenerf J, Betzler M. Comparing paper-based with electronic patient records: lessons learned during a study on diagnosis and procedure codes. J.Am.Med.Inform.Assoc. 2003;10(5):470–7.

**Stausberg et al. 2003d.** Stausberg J, Bromen K, Maier I, Kröger K, Niebel W. Integration von Informations- und Qualitätsmanagement: Statusbericht eines interdisziplinären Projekts zur Prävention und Behandlungsoptimierung des Dekubitus. Medizinische Informatik, Biometrie und Epidemiologie in Medizin und Biologie 2003;34:286–8.

**Stausberg et al. 2012.** Stausberg J, Pritzkuleit R, Schmidt CO, Schrader T, Nonnemacher M. Indicators of data quality: revision of a guideline for networked medical research. Stud Health Technol Inform. 2012;180:711–5.

**Steil et al. 2004.** Steil H, Amato C, Carioni C, Kirchgessner J, Marcelli D, Mitteregger A, Moscardo V, Orlandini G, Gatti E. EuCliD–a medical registry. Methods Inf.Med 2004;43(1):83–8.

**Stevens et al. 2008.** Stevens W, Stevens G, Kolbe J, Cox B. Comparison of New Zealand Cancer Registry data with an independent lung cancer audit. N Z Med J. 2008;121(1276):29–41.

**Storey/Wang 1998.** Storey VC, Wang RY. Modeling Quality Requirements in Conceptual Database Design. In: Proceedings of the 1998 Conference on Information Quality. Cambridge, MA; 1998: 64–87.

**Streit 2013.** Streit B. Der Study Advisor in OpEN.SC – Prozessorientierte Unterstützung des Entwurfs von klinischen Studien. Brandenburg, 2013.

**Streit et al. 2013.** Streit B, Hanss S, Zhou Y, Niepage S, Schrader T. Bewertung der Datenqualität des OpEN.SC Metadaten-Repositories mithilfe des Data Quality Assessment Frameworks. Braunschweig, 2012.

**Svolba 1999.** Svolba, G. Statistical quality control in clinical trials. (60). Dissertationen der Universität Wien. Wien: WUV-Universitätsverlag; 1999.

**Svolba/Bauer 1999.** Svolba G, Bauer P. Statistical quality control in clinical trials. Control Clin.Trials 1999;20(6):519–30.

**Swart 2004.** Swart E. Gute Praxis Sekundärdatenanalyse (GPS) – Adaptation der GEP an die Bedingungen der Sekundärdatenanalyse. Das Gesundheitswesen 2004;66:614.

**Swart et al. 2005.** Swart E, Ihle P, Geyer S, Grobe T, Hofmann W. GPS – Gute Praxis Sekundärdatenanalyse. Das Gesundheitswesen 2005;67:416–21.

**Swart/Ihle 2005.** Swart E, Ihle P (Herausgeber). Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. Bern: Verlag Hans Huber, Hogrefe AG; 2005.

**Taggart et al. 2012.** Taggart J, Liaw ST, Dennis S, Yu H, Rahimi A, Jalaludin B, Harris M. The University of NSW electronic practice based research network: disease registers, data quality and utility. Stud Health Technol Inform. 2012;178:219–27.

**Teperi 1993.** Teperi J. Multi method approach to the assessment of data quality in the Finnish Medical Birth Registry. J.Epidemiol.Community Health 1993;47(3):242–7.

**Teppo et al. 1994.** Teppo L, Pukkala E, Lehtonen M. Data quality and quality control of a population-based cancer registry. Experience in Finland. Acta Oncol. 1994;33(4):365–9.

**Thiru et al. 2005.** Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. BMJ 2003;326(7398):1070.

**Thoburn et al. 2007.** Thoburn KK, German RR, Lewis M, Nichols PJ, Ahmed F, Jackson-Thompson J. Case completeness and data accuracy in the Centers for Disease Control and Prevention's National Program of Cancer Registries. Cancer. 2007;109(8):1607–16.

**Tolonen et al. 2006.** Tolonen H, Dobson A, Kulathinal S; WHO MONICA Project. Assessing the quality of risk factor survey data: lessons from the WHO MONICA Project. Eur J Cardiovasc Prev Rehabil 2006;13(1):104–14.

**Topp et al. 1997.** Topp M, Langhoff-Roos J, Uldall P. Validation of a cerebral palsy register. J.Clin.Epidemiol. 1997;50(9):1017–23.

**Tuble 2011.** Tuble SC Perfusion Downunder Collaboration Database--data quality assurance: towards a high quality clinical database. J Extra Corpor Technol. 2011;43(1):P44–51.

**Tudur Smith et al. 2012.** Tudur Smith C, Stocken DD, Dunn J, Cox T, Ghaneh P, Cunningham D, Neoptolemos JP. The value of source data verification in a cancer clinical trial. PLoS One. 2012;7(12):e51623.

**Vantongelen et al. 1989.** Vantongelen K, Rotmensz N, van der Schueren E. Quality control of validity of data collected in clinical trials. EORTC Study Group on Data Management (SGDM). Eur.J.Cancer Clin.Oncol. 1989;25(8):1241–7.

**VdAK et al. o.J.** Verband der Angestellten-Krankenkassen, Arbeiter-Ersatzkassen-Verband, AOK-Bundesverband, Bundesverband der Betriebskrankenkassen, IKK Bundesverband, Bundesknappschaft, Bundesverband der landwirtschaftlichen Krankenkassen, See-Krankenkasse. Fehlbelegungsprüfung. Richtlinie nach § 282 SGB V. http://medinfoweb.de/apps/webeditor/files/rl_fehlbelegung.pdf [Zugriff am 21.04.2014].

**Venet et al. 2012.** Venet D, Doffagne E, Burzykowski T, Beckers F, Tellier Y, Genevois-Marlin E, Becker U, Bee V, Wilson V, Legrand C, Buyse M. A statistical approach to central monitoring of data quality in clinical trials. Clin Trials. 2012 Dec;9(6):705–13.

**Verhulst et al. 2012.** Verhulst K, Artiles-Carloni L, Beck M, Clarke JT, Neto JC, Cox GF, Fernhoff PM, Guffon N, Kong Y, Martins AM, Tylki-Szymanska A, Whitley CB, Wijburg FA, Wraith EJ, Koepper CM. Source document verification in the Mucopolysaccharidosis Type I Registry. Pharmacoepidemiol Drug Saf. 2012;21(7):749–752.

**Vestberg et al. 1997.** Vestberg K, Thulstrup AM, Sorensen HT, Ottesen P, Sabroe S, Vilstrup H. Data quality of administratively collected hospital discharge data for liver cirrhosis epidemiology. J.Med.Syst. 1997;21(1):11–20.

**Völzke et al. 2011.** Völzke H, Alte D, Schmidt CO, et al. Cohort profile: the study of health in Pomerania. Int J Epidemiol 2011;40:294–307.

**Wand/Wang 1996.** Wand Y, Wang RY. Anchoring data quality dimensions in ontological foundations. Commun. ACM 1996;39:86–95.

**Wang et al. 1993.** Wang RY, Kon HB, Madnick SE. Data quality requirements analysis and modeling. In: Proceedings of the 9th international conference on data engineering (ICDE), IEEE Computer Society Wien; 1993: 670–7.

**Ward et al. 2004.** Ward NS, Snyder JE, Ross S, Haze D, Levy MM. Comparison of a commercially available clinical information system with other methods of measuring critical care outcomes data. J Crit Care 2004;19(1):10–5.

**Weber 2013.** Weber GM. Federated queries of clinical data repositories: the sum of the parts does not equal the whole. J Am Med Inform Assoc 2013;20:e155–61.

**Weihs/Jessenberger 1999.** Weihs, C., Jessenberger, J. Statistische Methoden zur Qualitätssicherung und -optimierung in der Industrie. Weinheim, New York, Chichester, Brisbane, Singapore, Toronto: Wiley-VCH; 1999.

**Weiskopf/Weng 2013.** Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc. 2013;20(1):144–51.

**Weiss et al. 1993.** Weiss RB, Vogelzang NJ, Peterson BA, Panasci LC, Carpenter JT, Gavigan M, Sartell K, Frei E III, McIntyre OR. A successful system of scientific data audits for clinical trials. A report from the Cancer and Leukemia Group B. JAMA 1993;270(4):459–64.

**Weiss 1998.** Weiss RB. Systems of protocol review, quality assurance, and data audit. Cancer Chemother.Pharmacol. 1998;42 Suppl:S88-S92.

**Welzer et al. 2005.** Welzer T, Brumen B, Golob I, Sanchez JL, Druzovec M. Diagnostic process from the data quality point of view. J.Med.Syst. 2005;29(1):59–63.

**Whitney et al. 1998.** Whitney CW, Lind BK, Wahl PW. Quality assurance and quality control in longitudinal studies. Epidemiol Rev. 1998;20(1):71–80.

**Wichmann et al. 2012.** Wichmann HE, Kaaks R, Hoffmann W, Jockel KH, Greiser KH, Linseisen J. Die Nationale Kohorte. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 2012;55:781–7.

**Williams 1997.** Williams GW. Errors, misunderstandings, and fraud. Control Clin.Trials 1997;18(6):661–6.

**Winter et al. 2003.** Winter M, Herrmann C, Helfert M. Datenqualitätsmanagement für Data-Warehouse-Systeme – Technische und organisatorische Realisierung am Besipiel der Credit Suisse. In: Maur E, Winter R, editors. Data Warehouse Management: Das St. Galler Konzept zur ganzheitlichen Gestaltung der Informationslogistik. Heidelberg: Springer; 2003: 221–40.

**Witte et al. 2000.** Witte PU, Schenk J, Schwarz JA, Kori-Lindner C. Ordnungsgemässe klinische Prüfung. Good Clinical Practice. 5.ed. Berlin: E. Habrick Verlag; 2000.

**Wood/Youatt 1979.** Wood W, Youatt R. A computer feedback system for clinical research. Comput.Programs Biomed. 1979;9(1):80–94.

**Wu et al. 2008.** Wu Y, Takkenberg JJ, Grunkemeier GL. Measuring follow-up completeness. Ann Thorac Surg. 2008;85(4):1155–7.

**Xian et al. 2012.** Xian Y, Fonarow GC, Reeves MJ, Webb LE, Blevins J, Demyanenko VS, Zhao X, Olson DM, Hernandez AF, Peterson ED, Schwamm LH, Smith EE. Data quality in the American Heart Association Get With The Guidelines-Stroke (GWTG-Stroke): results from a national data validation audit. Am Heart J. 2012;163(3):392–8, 398.

**Zettelmeyer et al. 2005.** Zettelmeyer U, Pfistner B, Diehl V, Paulus, U. On-site Monitoring als Instrument zur Qualitätssicherung von Daten bei klinischen Studien. Düsseldorf, Köln: German Medical Science; 2005. http://www.egms.de/en/meetings/gmds2005/05gmds317.shtml [Zugriff am 07.04.2014].

**Ziegler et al. 2011.** Ziegler A, Antes G, König IR. Bevorzugte Report Items für systematische Übersichten und Meta-Analysen: Das PRISMA-Statement. Dtsch Med Wochenschr 2011;136:e9–e15.

**Zorn/Ollenschläger 1999.** Zorn U, Ollenschläger G. Qualitätsbestimmung in der medizinischen Versorgung – ein universelles Entwicklungsschema für Qualitätsindikatoren. Z.ärztl.Fortb.Qual.sich. 1999;93:123–8.

## Abkürzungsverzeichnis

| | |
|---|---|
| ABPI | Association of the British Pharmaceutical Industry |
| ADR | Adverse Drug Reaction |
| AG | Arbeitsgruppe |
| AGENS | Arbeitsgruppe Erhebung und Nutzung von Sekundärdaten |
| AGREE | Appraisal of Guidelines for Research and Evaluation in Europe |
| AIDS | acquired immunodeficiency syndrome |
| AHRQ | Agency for Healthcare Research and Quality |
| AHIMA | American Health Information Management Association |
| AIDS | Acquired Immune Deficiency Syndrome |
| AkdÄ | Arzneimittelkommission der deutschen Ärzteschaft |
| AMG | Arzneimittelgesetz |
| ANSI | American National Standards Institute |
| anuba | Aufbau und Nutzung von Bildungsnetzwerken zur Entwicklung und Erprobung von Ausbildungsmodulen in IT- und Medienberufen |
| APACHE | Acute Physiology and Chronic Health Evaluation |
| AWMF | Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften |
| ÄZQ | Ärztliches Zentrum für Qualität in der Medizin |
| BÄK | Bundesärztekammer |
| BMGS | Bundesministerium für Gesundheit und Soziale Sicherung |
| BMI | Body Mass Index |
| BPM | Business Process Modeling |
| BQS | Bundesgeschäftsstelle Qualitätssicherung |
| BzgA | Bundeszentrale für gesundheitliche Aufklärung |
| CALGB | Cancer and Leukemia Group B |
| CAP | Community Acquired Pneumonia |
| CARF | Commission on Accreditation of Rehabilitation Facilities |
| CDISC | Clinical Data Interchange Standards Consortium |
| CDISC-PR | CDISC Protocol Elements for Regulated Clinical Trials |
| CDSS | Clinical Decision Support System |
| CEO | Chief Executive Officer |
| CFR | Code of Federal Regulations |
| CIBMTR | Centre for International Blood and Marrow Transplant Research |
| CIO | Chief Information Officer |
| CIS | Carcinoma in situ oder Clinical Information System |
| COPD | Chronic Obstructive Pulmonary Disease |

| | |
|---|---|
| CP | Cerebral Palsy |
| CPR | Computer-based Patient Record |
| CQI | Continuous Quality Improvement |
| CRF | Case Report Form |
| CPCRA | Terry Beirn Community Programs for Clinical Research on AIDS |
| CSS | Cancer Surveillance System |
| CTD | Common Technical Document |
| CTN | Clinical Trials Network |
| CTQ | Critical to Quality |
| DAC | Data Audit Committee |
| DAE | Deutsche Arbeitsgemeinschaft für Epidemiologie |
| DB | Datenbank |
| DBMS | Datenbank Management System |
| DCCG | Danish Colorectal Cancer Group |
| DCO | Death Certificate Only |
| DEGEMED | Die Deutsche Gesellschaft für Medizinische Rehabilitation |
| DELBI | Deutsches Leitlinien-Bewertungsinstrument |
| DGSMP | Deutsche Gesellschaft für Sozialmedizin und Prävention |
| DICOM | Digital Imaging and Communications in Medicine |
| DIN | Deutsches Institut für Normung |
| DMC | Data Monitoring Committee |
| DNPR | Danish National Patient Registry |
| DPMO | Defects per Million Opportunities |
| DQM | Datenqualitätsmanagement (Data Quality Management) |
| DQMS | Data Quality Monitoring System |
| DQP | Data Quality Probe |
| DQV | Datenqualitätsvektor |
| DRST | Deutsche Register für Stammzelltransplantation |
| DV | Datenverarbeitung |
| DWQ | Foundations of Data Warehouse Quality |
| EACTS | European Association for Cardio-Thoracic Surgery |
| EBM | Evidence Based Medicine |
| EBMT | European Group for Blood and Marrow Transplantation |
| eBPRN | Electronic Practice Based Research Network |
| EDC | Electronic Data Capture |
| EDQM | European Department for Quality of Medicines |

| | |
|---|---|
| EFQM | European Foundation for Quality Management |
| EDV | Elektronische Datenverarbeitung |
| EHR | Electronic Health Record |
| EKG | Elektrokardiogramm |
| EMEA | European Agency for the Evaluation of Medicinal Products |
| EN | Euronorm |
| EORTC | European Organisation for Research and Treatment of Cancer |
| EPA | Elektronische Patientenakte |
| ESP | Estimated survival probability |
| ESPRIT | European Strategic Program of Research and development in Information Technology |
| ESTS | European Society of Thoracic Surgeons |
| EU | Europäische Union |
| EUROCAT | European Surveillance of Congenital Anomalies |
| FDA | Food and Drug Administration |
| FAQ | Frequently Asked Questions |
| F & E | Forschung und Entwicklung |
| FMEA | Failure Mode and Effects Analysis |
| FQS | Forschungsstelle Qualitätssicherung |
| FTA | Failure Tree Analysis |
| FTP | File Transfer Protocol |
| FU | Follow-up |
| GCP | Good Clinical Practice |
| GDP | Good Distribution Practice |
| GEP | Gute Epidemiologische Praxis |
| GKV | Gesetzliche Krankenversicherung |
| GLIF | Guideline Interchange Format |
| GLP | Good Laboratory Practice |
| GMDS | Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie |
| GMP | Good Manufacturing Practice |
| GQM | Goal-Question-Metrik |
| GQMG | Gesellschaft für Qualitätsmanagement in der Gesundheitsversorgung |
| GPS | Good Practice of Secondary Data Analysis |
| GR & R | Gage Repeatability & Reproducibility |
| GWTG | Get With The Guidelines |
| HEDIS | Health Plan Employer Data and Information Set |
| HIV | human immunodeficiency virus |

| | |
|---|---|
| HL7 | Health Level 7 |
| HTA | Health Technology Assessment |
| IARC | International Agency for Research on Cancer |
| IBE | Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie |
| ICC | Intraclass Correlation (Intraklassenkorrelation) |
| ICD | International Classification of Diseases |
| ICD-10 | Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision |
| ICH | International Conference on Harmonisation |
| ICNARC | Intensive Care National Audit & Research Centre |
| ID | Identifikator |
| IDDM | Insulinabhängiger Diabetes mellitus |
| IMIBE | Institut für Medizinische Informatik, Biometrie und Epidemiologie |
| IQWIG | Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen |
| ISO | International Organization for Standardization |
| IT | Informationstechnik |
| ITR | Illinois Trauma Registry |
| ITT | Intention to Treat |
| JCAHO | Joint Commission for the Accreditation of Healthcare Organizations |
| JCIA | Joint Commission International Accreditation |
| KAS | Klinisches Arbeitsplatzsystem |
| KBV | Kassenärztliche Bundesvereinigung |
| KHK | Koronare Herzkrankheit |
| KIS | Krankenhausinformationssystem |
| KKS | Koordinierungszentrum für Klinische Studien |
| KTQ | Kooperation für Transparenz und Qualität im Gesundheitswesen |
| LDS | Latter Day Saints |
| LOINC | Logical Observation Identifiers Names and Codes |
| MADRS | Montgomery-Åsberg Depression Rating Scale |
| MCO | Managed Care Organizations |
| MDK | Medizinischer Dienst der Krankenversicherung |
| MDPH | Michigan Department of Public Health |
| MDS | Medizinischer Dienst der Spitzenverbände der Krankenkassen |
| MeSH | Medical Subject Headings |
| MONICA | Monitoring of Trends and Determinants in Cardiovascular Disease |
| MRT | Magnetresonanztomographie |

| | |
|---|---|
| NCHS | National Center for Health Statistics |
| NCI | National Cancer Institute |
| NCQA | National Committee for Quality Assurance |
| NICE | National Intensive Care Evaluation |
| NIDA | National Institute of Drug Abuse |
| NIH | National Institute of Health |
| NHLBI | National Heart, Lung and Blood Institute |
| NHS | National Health Service |
| NPCR | National Program of Cancer Registries |
| NSAID | Nonsteroidal Antiinflammatory Drugs |
| NSAR | Nichtsteroidale Antirheumatika |
| NSABBP | National Surgical Adjuvant Breast and Bowel Project |
| NZCR | New Zealand Cancer Registry |
| ODM | Operational Data Model |
| o.J. | ohne Jahr |
| ONS | Office for National Statistics |
| OP | Operation |
| OpEN.SC | Open European Nephrology Science Center |
| OPS | Operationen- und Prozedurenschlüssel |
| OR | Odds Ratio |
| OZ II | Operatives Zentrum II |
| PAC | Performance Assessment Committee |
| PC | Personal Computer |
| PCDQ | Primary Care Data Quality |
| PDCA | Plan – Do – Check – Act |
| PDF | Portable Document Format |
| PDUC | Perfusion Downunder Collaboration |
| PFS | Progression-Free Survival (folgenloses Überleben) |
| PMTCT | Prevention of Mother-To-Child Transmission |
| POG | Pediatric Oncology Group |
| PSI | Pneumonia Severity Index |
| QC | Quality Control |
| QFD | Quality Function Deployment |
| QDX | Quality Data Exchange |
| QI | Quality Improvement |
| QI | Qualitätsindikator |

| | |
|---|---|
| QS | Qualitätssicherung |
| QTR | Queensland Trauma Registry |
| RAND | RAND Corporation, Kalifornien |
| RCA | Root Cause Analysis |
| RDE | Remote Data Entry |
| RECIST | Response Evaluation Criteria in Solid Tumors |
| RPZ | Risikoprioritätszahl |
| RR | Relatives Risiko |
| SAE | Serious Adverse Event (schwerwiegendes unerwünschtes Ereignis) |
| SAPS | Simplified Acute Physiology Score |
| SAS | Statistical Analysis System |
| SDTM | Study Data Tabulation Model |
| SDV | Source Data Verification |
| SGB | Sozialgesetzbuch |
| SHIP | Study of Health in Pomerania |
| SNOMED | Systematized Nomenclature of Human and Veterinary Medicine |
| SOA | Serviceorientierte Architektur |
| SOP | Standard Operating Procedure |
| SPSS | Statistical Package for the Social Sciences |
| SQL | Structured Query Language |
| TDWI | The Data Warehouse Institute |
| TMF | TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. |
| TPFN | True Positive to False Negative Ratio |
| TQM | Total Quality Management |
| UK | United Kingdom (Vereinigtes Königreich Großbritannien und Nordirland) |
| UML | Unified Modeling Language |
| USA | United States of America (Vereinigte Staaten von Amerika) |
| UTI | Urinary Tract Infection |
| VdAK | Verband der Angestellten-Krankenkassen |
| VDI | Verein Deutscher Ingenieure |
| WHO | World Health Organization |
| WSI | Whole Slide Images |
| XML | eXtensible Markup Language |

# Tabellen und Abbildungen

## Verzeichnis der Tabellen



## Verzeichnis der Abbildungen

# XIII Anhänge

## Anhang 1 – Kurzbeschreibung der statistischen Tabellen

Die statistischen Tabellen ermöglichen es, für häufige Konstellationen Konfidenzintervalle und Fallzahlen nachzuschlagen. Teil A enthält 95%-Konfidenzintervalle für Raten der Qualitätsindikatoren. Folgende Werte wurden hierbei für den Nenner (**n**) berücksichtigt: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000 und 10.000. Werte für den Zähler (**m**) wurden für alle **n** über die Raten zwischen 0 und 0,5 in Schritten von 0,01 ermittelt. Aus Teil B lässt sich die adjustierte Fallzahl für die Source Data Verification ablesen. Für die Rate **p** wurden dort in Schritten von 0,01 alle Werte von 0,01 bis 0,5 sowie als untere Grenze 0,0062 berücksichtigt, für die halbe Breite des 95%-Konfidenzintervalls (**δ**) die Werte 0,01, 0,02, 0,03, 0,04 und 0,05. Für die Anzahl von Beobachtungseinheiten (**N**) wurden folgende Werte angesetzt: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1.000, 2.000, 3.000, 4.000, 5.000, 6.000, 7.000, 8.000, 9.000 und 10.000. Die sich ergebenden Werte für die adjustierte Fallzahl (**n**) wurden aufgerundet.

Die statistischen Tabellen können bei der TMF-Geschäftsstelle angefordert werden bzw. stehen auf der Website der TMF zum Download bereit (info@tmf-ev.de/www.tmf-ev.de/Produkte/P020021).

# Anhang 2 – Literatursichtung 2005 bis 2013

23. März 2014

## 1 Einleitung

Für Version 1.0 der Leitlinie wurde eine umfassende Literatursichtung vorgenommen. Diese umfasste thematisch neben der Datenqualität auch weitere relevante Themengebiete wie die Erstellung von Leitlinien, die Methologie von Qualitätsindikatoren sowie statistische und andere Methoden des Qualitätsmanagements. Zur Quellensuche wurden Literaturdatenbanken, allgemeine Suchmaschinen sowie Homepages einschlägiger Institutionen verwendet. Eingeschlossen wurden Quellen bis Frühjahr 2006. Die Zusammenfassung von Literatursichtung und Expertenbefragung bis 2005/2006 steht unter www.tmf-ev.de/Leitlinie-Datenqualitaet zum Download zur Verfügung.

Die hier zusammengefasste Literatursichtung ergänzt die Arbeiten aus Version 1.0 der Leitlinie um den Zeitraum 2005 bis Anfang 2013. Hierbei wurde eine Konzentration auf die Kernthemen der Leitlinie vorgenommen. So wurde als Literaturdatenbank ausschließlich Medline genutzt sowie alleinig die Suchanfrage zur Recherche nach Erkenntnissen zu Indikatoren zur Datenqualität, zum Orginaldatenabgleich sowie zum Feedback zur Datenqualität eingesetzt. Ziel war es, bei diesen Kernthemen dem Vorgehen der Literaturrecherche in Version 1.0 der Leitlinie zu entsprechen. In Version 2.0 hat sich allerdings die Beziehung zwischen Literatursichtung und Empfehlungen geändert. Diese stehen nun komplementär zueinander. Die Weiterentwicklung der Empfehlungen setzt im Wesentlichen die Ergebnisse aus dem vorangehenden Projekt mit Ergänzung von Qualitätsindikatoren um (s. Einleitung). Die Literatursichtung wurde parallel mit der Einarbeitung dieser Ergebnisse vorgenommen. Die Weiterentwicklung spiegelt daher nicht zwingend die bestmögliche Evidenz aus der hier dargestellten Literatursichtung wider, sondern die Erkenntnisse des vorangehenden Projektes. Dennoch wurde auf einen Abgleich geachtet, allerdings ohne neue Qualitätsindikatoren aufzunehmen, Qualitätsindikatoren zu verwerfen oder Qualitätsindikatoren in ihrer Kernaussage zu verändern.

## 2 Literaturrecherche

### 2.1 Einführung

Recherchiert wurde am 6. März 2013 in der Literaturdatenbank Medline über das Portal PubMed (http://www.pubmed.org/). Damit ist prinzipiell Literatur seit 1946 eingeschlossen. Die Suche wurde eingeschränkt auf Treffer seit dem 1.1.2005. Bei der Verwendung von „related citations“ wurden Einträge vor 2005 manuell entfernt.



## 2.2 Suchanfragen

data quality[Title] AND ("2005/01/01"[Date – Publication] : "3000"[Date – Publication]), 284 Treffer

"registries"[TIAB] NOT Medline[SB] OR "registries"[MeSH Terms] OR registry[-Text Word] AND "data quality"[All Fields] AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 191 Treffer

(("registries"[TIAB] NOT Medline[SB]) OR "registries"[MeSH Terms] OR registry[Text Word]) AND ("quality assessment"[All Fields] OR "quality control"[All Fields]) AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 338 Treffer

cohort[All Fields] AND ("quality assessment"[All Fields] OR "quality control"[All Fields]) AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 534 Treffer

Related citations zu Gissler M, Teperi J, Hemminki E, Merilainen J. Data quality after restructuring a national medical registry. Scand J Soc Med. 1995 Mar;23(1):75–80. PMID: 7784857 [PubMed – indexed for MEDLINE], 34 Treffer

Related citations zu Welzer T, Brumen B, Golob I, Sanchez JL, Druzovec M. Diagnostic process from the data quality point of view. J Med Syst. 2005 Feb;29(1):59–63. PMID: 15839332 [PubMed – indexed for MEDLINE], 84 Treffer

Related citations zu Crerand WJ, Lamb J, Rulon V, Karal B, Mardekian J. Building data quality into clinical trials. J AHIMA. 2002 Nov-Dec;73(10):44–6, 48–53, 2; quiz 55–6. PMID: 12432815 [PubMed – indexed for MEDLINE], 84 Treffer

Related citations zu Brown PJ, Warmington V. Data quality probes-exploiting and improving the quality of electronic patient record data and patient care. Int J Med Inform. 2002 Dec 18;68(1–3):91–8. PMID: 12467794 [PubMed – indexed for MEDLINE], 93 Treffer

("fraud"[MeSH Terms] OR "fraud" [Text Word]) AND ("clinical trial"[Publication Type] OR " clinical trials as topic"[MeSH Terms] OR " clinical trial"[Text Word]) AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 61 Treffer

"data collection"[MeSH Terms] AND "medical registry" [All Fields] AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 14 Treffer

"data accuracy"[All Fields] AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 148 Treffer

Arts D[Author] AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 9 Treffer

Knatterud GL[Author] AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 8 Treffer

Ollenschlager G[Author] AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 37 Treffer

"source data verification"[All Fields] AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 16 Treffer

"feedback"[MeSH Terms] AND ("medical registry"[All Fields] OR "cohort" [All Fields]) AND "2005/01/01"[Date – Publication] : "3000"[Date – Publication], 132 Treffer

## 2.3 Literaturauswahl

Um doppelte Literaturstellen bereinigt ergab die Recherche 1916 Treffer. Abbildung 11 zeigt das Vorgehen und die Ergebnisse der Literaturauswahl in der Übersicht. Für eine Aufnahme in die Leitlinie wurde inhaltlich eine Auseinandersetzung mit Indikatoren von Datenqualität, mit einem Originaldatenabgleich oder mit einem Feedback von Datenqualität an Erhebungszentren gefordert. Eine reine Anwendung dieser Aspekte war für einen Einschluss nicht ausreichend. Als Bewertungskategorien standen „zutreffend", „nicht zutreffend" und „unklar" zur Verfügung.

Zur Entwicklung eines gemeinsamen Verständnisses über diese Bewertungskriterien wurde initial eine Teilmenge von 100 der 1.916 Literaturstellen an Hand der Abstracts durch alle drei Autoren gesichtet. Abweichungen wurden gemeinsam besprochen und eine Entscheidung zuerst nach Majorität und bei kompletter Abweichung im Konsens getroffen. Von den verbleibenden 1.816 Literaturstellen wurden 100 verblindet zur Reliabilitätsprüfung durch alle drei Autoren und 1.716 von jeweils einem Autor gesichtet. Bei der Stichprobe zur Reliabilitätsprüfung wurde die Entscheidung wieder zuerst nach Majorität und bei kompletter Abweichung im Konsens getroffen. Unklarheiten bei den nur durch einen Autor gesichteten Literaturstellen wurden im Konsens aufgelöst. Volltexte wurden von denjenigen Literaturstellen beschafft, die mit zutreffend oder unklar bewertet waren. Als Reliabilitätsmaß wurde das Kappa nach Siegel/Castellan (1988) mit dem Script MKAPPASC.SPS mit IBM SPSS Statistics 21 berechnet. Kappa lag in der ersten Stichprobe bei 0,42 und in der zweiten bei 0,39. Nach Landis/Koch (1977) entspricht dies einer moderaten bzw. ordentlichen Übereinstimmung.

Nach Ausschluss von 3 Literaturstellen mit einer anderen Sprache als Deutsch oder Englisch wurden die verbleibenden 107 Literaturstellen an Hand der Volltexte in weiteres Mal bewertet. Es kamen die gleichen Bewertungskategorien zum Einsatz. Zusätzlich war bei einer Bewertung mit „zutreffend" der zur positiven Bewertung führende inhaltliche Aspekt anzugeben. Die 107 Literaturstellen wurden von JS komplett sowie mit einer Aufteilung von 53 und 54 Literaturstellen von DN und MN gesichtet. Abweichungen wurden bilateral aufgelöst und eine definitive Entscheidung über den Einschluss in die Leitlinie getroffen.




Als Ergebnis der Literaturauswahl werden 39 Literaturstellen ausführlich in der Leitlinie dargestellt.

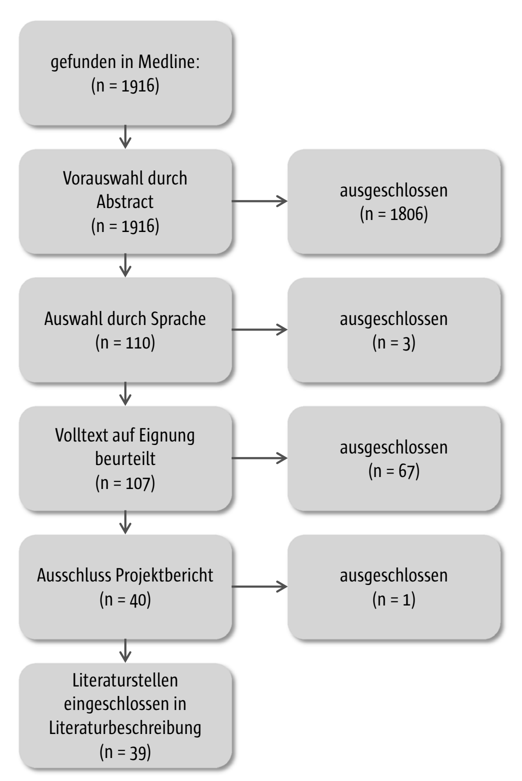



Abb. 11 Vorgehen und Ergebnisse der Literaturauswahl, dargestellt gemäß dem PRISMA-Statement (Ziegler et al. 2011)

# 3 Ergebnisse

## 3.1 Struktur

Die Beschreibungen wurden durch die Autoren DN und MN verfasst. Die Darstellung erfolgt strukturiert mit folgenden Rubriken:

1. Art des wissenschaftlichen Vorhabens: <Analyse von Datenqualität|Intervention zur Verbesserung der Datenqualität|Health Technology Assessment (Nutzenbewertung von Datenqualität)|systematisches Review|Konzeptvorstellung>
2. Herkunft der Autoren: Land
3. Bezug zum Thema: <Qualitätsindikatoren|Feedback|SDV>, Mehrfachangabe möglich

4. Ort und Art der Intervention: Angabe nur bei Interventionsstudien oder einem HTA
5. Datenbestand
6. Qualitätsindikatoren: Benennung nach den Autoren
7. Schlussfolgerungen der Autoren: Darstellung der von den Autoren der Literaturstelle selbst gezogenen Schlussfolgerungen.
8. Zusammenfassung der Publikation
9. Kommentar

## 3.2 Übersetzung englischer Bezeichnungen

Englische Bezeichnungen für qualitätsrelevante Aspekte wurden wie folgt übersetzt: Aktualität (currency), angemessenes Datenvolumen (appropriate amount of data), Brauchbarkeit (usefulness of data items), Definition (definition), Duplizität (prevention of duplicates), Empfindlichkeit (responsiveness of data items), Erreichbarkeit (accessibility), Genauigkeit (accuracy), Generalisierbarkeit (generalizability), Glaubwürdigkeit (believability), Granularität (granularity), Inkonsistenz (inconsistency), Konkordanz (concordance), Kontextannotation (contextualization), Korrektheit (correctness), Objektivität (objectivity), Plausibilität (plausibility), politische Relevanz (policy relevance), Präzision (precision), räumliche Stabilität (spatial stability), Rekrutierungsrate (rate of enrolment), Relevanz (relevancy), Übereinstimmung (agreement), Unrichtigkeit (incorrectness), Unvollständigkeit (incompleteness), Validität (validity), Verfügbarkeit (availability), Vergleichbarkeit (comparability), Vollständigkeit (comprehensiveness), Vollzähligkeit (completeness), Vorhersagewert (predictive value), Widerspruchsfreiheit (consistency), Zeitnähe (timeliness), Zuverlässigkeit (reliability).

## 3.3 Übersicht

Tabelle 14 zeigt die 39 Literaturstellen mit den Angaben zu den Rubriken 1 bis 3 und 6 in der Übersicht. Der Kurzverweis in Tabelle 14 sowie die Überschriften in Kapitel 3.4 dieses Anhangs entsprechen der Notation im Literaturverzeichnis. Dort sind auch die Literaturstellen aufgeführt.

Tab. 14 Ausgewählte Eigenschaften der Literaturstellen

| Literaturstelle | Art | Herkunft | Bezug | Qualitätsindikatoren |
|---|---|---|---|---|
| Baigent et al. 2008 | Konzept | UK | QI, SDV | nicht anwendbar |
| Berner et al. 2005 | HTA | USA | QI | Genauigkeit, Vollständigkeit |
| Botsis et al. 2010 | Analyse | Norwegen, USA | QI, SDV | Inkonsistenz, Unrichtigkeit, Unvollständigkeit |



| Literaturstelle | Art | Herkunft | Bezug | Qualitätsindikatoren |
|---|---|---|---|---|
| Bray/Parkin 2009 | Review | Norwegen | QI | Aktualität, Validität, Vergleichbarkeit |
| Brender et al. 2008 | Analyse | USA | SDV | Übereinstimmung |
| Bronnert et al. 2012 | Konzept | USA | QI | Aktualität, Definition, Erreichbarkeit, Genauigkeit, Granularität, Präzision, Relevanz, Vollständigkeit, Widerspruchsfreiheit, Zeitnähe |
| Brower et al. 2006 | Review | Niederlande | Feedback, QI | Korrektheit, Vollständigkeit |
| Chiba et al. 2012 | Analyse | Japan | QI | Genauigkeit, Relevanz, Vollständigkeit |
| Choquet et al. 2010 | Konzept | Frankreich | QI | nicht anwendbar |
| Couchoud et al. 2013 | Konzept | Frankreich | QI | Aktualität, Brauchbarkeit, Duplizität, Empfindlichkeit, Genauigkeit, Validität, Vergleichbarkeit, Vollständigkeit, Vollzähligkeit |
| De 2011 | Konzept | Indien | SDV | nicht anwendbar |
| Duda et al. 2012 | Analyse | USA | SDV | Korrektheit, Vollständigkeit |
| Dyck et al. 2007 | Konzept | USA | QI | Validität, Vollständigkeit, Zuverlässigkeit |
| França et al. 2008 | Analyse | Brasilien | QI | Generalisierbarkeit, Korrektheit, politische Relevanz, Zuverlässigkeit |
| Kahn et al. 2012 | Konzept | USA | QI | angemessenes Datenvolumen, Glaubwürdigkeit, Korrektheit, Objektivität, Zeitnähe |
| Krzych et al. 2011 | Analyse | UK, Polen | QI | Aktualität, Rekrutierungsrate, Vollständigkeit |
| Larsen et al. 2009 | Analyse | Norwegen | QI | Aktualität, Korrektheit, Vergleichbarkeit, Vollständigkeit |
| Loane et al. 2011 | Konzept | UK | QI | Korrektheit, Zeitnähe, Verfügbarkeit, Vollständigkeit, Vollzähligkeit |
| Macefield et al. 2013 | Review | UK | SDV | nicht anwendbar |
| Maruszewski et al. 2005 | Analyse | Multinational | QI, SDV | Genauigkeit, Vollständigkeit |
| McKenzie et al. 2005 | Analyse | Australien | QI | Konkordanz |
| Messenger et al. 2012 | Analyse | USA | Feedback, QI, SDV | Korrektheit, Vollständigkeit, Widerspruchsfreiheit |

| Literaturstelle | Art | Herkunft | Bezug | Qualitätsindikatoren |
|---|---|---|---|---|
| Mphatswe et al. 2012 | Intervention | Südafrika | Feedback, QI, SDV | Genauigkeit, Vollständigkeit |
| Nahm et al. 2008 | Intervention | USA | SDV | nicht anwendbar |
| Sáez et al. 2012 | Konzept | Spanien | QI | Duplizität, Kontextannotation, Korrektheit, räumliche Stabilität Vollständigkeit, Vorhersagewert, Widerspruchsfreiheit, Zeitnähe, Zuverlässigkeit |
| Salati et al. 2011 | Konzept | Italien | QI | Glaubwürdigkeit, Korrektheit/Genauigkeit, Vollständigkeit, Widerspruchsfreiheit |
| Shabestari/Roudsari 2013 | Konzept | Kanada | QI | Konkordanz, Korrektheit, Plausibilität, Vollständigkeit, Zeitnähe |
| Sigurdardottir et al. 2012 | Analyse | Island | QI, SDV | Aktualität, Validität, Vollzähligkeit |
| Stevens et al. 2008 | Analyse | Neuseeland | QI | Genauigkeit, Vollständigkeit |
| Taggart et al. 2012 | Analyse | Australien | Feedback, QI | Korrektheit, Vollständigkeit, Widerspruchsfreiheit |
| Thoburn et al. 2007 | Analyse | USA | QI | Korrektheit, Vollständigkeit |
| Tolonen et al. 2006 | Konzept | Finnland | QI | nicht anwendbar |
| Tuble 2011 | Analyse | Australien | QI, SDV | Genauigkeit, Vollständigkeit |
| Tudur Smith et al. 2012 | HTA | UK | SDV | nicht anwendbar |
| Venet et al. 2012 | Konzept | Belgien | QI | nicht anwendbar |
| Verhulst et al. 2012 | Analyse | Niederlande | SDV | Korrektheit |
| Weiskopf/Weng 2013 | Review | USA | QI | Konkordanz, Korrektheit, Plausibilität, Vollständigkeit, Zeitnähe |
| Wu et al. 2008 | Konzept | USA | QI | Rekrutierungsrate, Vollzähligkeit |
| Xian et al. 2012 | Analyse | USA | QI | Genauigkeit, Zuverlässigkeit. |

HTA = Health Technology Assessment, QI = Qualitätsindikatoren, SDV = Source Data Verification (Originaldatenabgleich), UK = United Kingdom (Vereinigtes Königreich Großbritannien und Nordirland), USA = United States of America (Vereinigte Staaten von Amerika)




## 3.4 Literaturstellen

### Baigent et al. 2008

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Vereinigtes Königreich
3. Bezug zum Thema: Qualitätsindikatoren, SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: nicht anwendbar
7. Schlussfolgerungen der Autoren
   Die Qualitätssicherung in Studien muss nicht mit einem intensiven Vor-Ort-Monitoring verbunden sein. Eine Risikoabschätzung bezüglich möglicher Fehler sollte die Basis für die Entwicklung eines optimalen Monitoringplans bilden. So lassen sich die vorhandenen Ressourcen bestmöglich nutzen.
8. Zusammenfassung der Publikation
   Der Artikel fasst die Diskussionen zusammen, die im Jahre 2007 auf dem Sensible Guidelines Workshop in den USA geführt wurden. Es geht um die Frage, wie verschiedene Arten des Monitorings bei randomisierten Studien die Datenqualität verbessern können. In randomisierten Studien sind vier verschiedene Arten von Fehlern möglich: Fehler im Studiendesign (z.B. ungenügende Fallzahl), fehlerhafte Durchführung der Studie (z.B. Abweichung vom Studienprotokoll), fehlerhafte Daten (zufällig oder durch Betrug entstanden) sowie Fehler bei der statistischen Analyse der Studiendaten und der Interpretation der Studienergebnisse. Dem stehen verschiedene Formen des Monitorings gegenüber: Kontrolle durch ein Datenmonitoring-Komitee als Teil eines Studienkommittees, zentrales Monitoring durch die Datenmanagementzentrale und ein Vor-Ort-Monitoring in den einzelnen Studienzentren. Die Autoren ordnen diese Formen den Fehlerarten zu, für deren Entdeckung sie nach Meinung der Autoren besonders geeignet sind. Das zentrale Monitoring erlaubt die Betrachtung von Qualitätsindikatoren (z.B. zeitlicher Verzug bei der Datenerfassung). Die Autoren schlagen vor, für randomisierte Studien jeweils eine Risikoanalyse bezüglich der o.g. Fehlerarten durchzuführen und einen dazu passenden Monitoringplan zu entwerfen.
9. Kommentar
   Die Autoren postulieren, ähnlich wie es die vorliegende Leitlinie tut, eine Ausrichtung des Monitorings auf eine möglichst effiziente Nutzung der vorhandenen Ressourcen.

### Berner et al. 2005

1. Art des wissenschaftlichen Vorhabens: Health Technology Assessment
2. Herkunft der Autoren: USA

3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: Untersuchung von Genauigkeit und Vollständigkeit bei der Erfassung von Patienteninformationen in niedergelassenen Arztpraxen und deren Auswirkung auf Clinical Decision Support Systems (CDSS).
5. Datenbestand: 178 Patientenakten simulierter Patienten.
6. Qualitätsindikatoren: Genauigkeit, Vollständigkeit
7. Schlussfolgerungen der Autoren
   Clinical Decision Support Systems (CDSS) können die Patientensicherheit im Gesundheitswesen durch Empfehlungen zu Verschreibungen von potenziell gefährlichen Medikamenten verbessern. Diese CDSS benötigen allerdings hochqualitative Daten. Die Studie zeigt, dass CDSS im Bereich der niedergelassenen Arztpraxen auf Grund mangelnder Datenvollständigkeit und Genauigkeit keine zuverlässigen Entscheidungen liefern.
8. Zusammenfassung der Publikation
   Vorhergehende Arbeiten haben gezeigt, dass sich die Verwendung von CDSS im stationären Umfeld positiv auf die Verringerung von Behandlungsfehlern auswirken. Dies geschieht durch das Auslösen von Alarmsignalen bei fehlerhafter medizinischer Intervention basierend auf Daten der Patientenakte. In dieser Arbeit wird speziell auf die Auswirkungen von Genauigkeit und Vollständigkeit der Daten medizinischer Akten auf ein CDSS, das eine Risikoabschätzung zu Magendarmblutungen bei der Verschreibung von Nonsteroidal Antiinflammatory Drugs (NSAIDs) liefert, eingegangen. Als Eingabedaten wurden Patientenakten von standardisierten Patienten verwendet. Die standardisierten Patienten wurden darauf trainiert, verschiedene Muskel-Skelett-Beschwerden vorzutragen. Dies sind Krankheiten, bei denen eine NSAID-Verschreibung sinnvoll ist. Zudem fällt jeder standardisierten Patient durch Faktoren, die sich aus dem Alter oder der simulierten Krankengeschichte ergeben, in die Gruppe der Hochrisikopatienten für interne Blutungen des Magen-Darm-Traktes bei NSAID-Verschreibung. Bei korrekter Aufnahme aller Risikofaktoren und Übermittlung der Daten an das CDSS würde ein Alarm, der vor einer NSAID-Verabreichung warnt, ausgelöst werden. Die standardisierten Patienten wurden zu 60 niedergelassenen Ärzten geschickt und in einem Aufnahmegespräch befragt. Die diesbezüglichen Aufzeichnungen wurden von den Arztpraxen an das Studienteam weitergeleitet und insgesamt 178 Patientenakten anhand der in den Aufzeichnungen enthaltenen Informationen erstellt. Dabei erfolgte eine Überprüfung, ob zu den Risikofaktoren Informationen vorhanden waren und falls ja, ob diese auch korrekt ermittelt wurden. Hierdurch konnten Genauigkeit und Vollständigkeit der resultierenden Patientenakten abgeleitet werden. Lediglich 30–40% aller Risikofaktoren wurden aufgedeckt. Die erfassten Daten waren mit 91–96% jedoch weitestgehend korrekt. Das CDSS erkannte aufgrund der Datenmängel einen Hochrisikopatienten nur zu 22,5%.




9. Kommentar
   Die Publikation verdeutlicht, wie sich Probleme der Datenqualität unmittelbar negativ auf den Patienten auswirken können. Im gegebenen Beispiel wurden 77,5% der simulierten Hochrisikopatienten durch das CDSS nicht erkannt. Eine Verschreibung von NSAID hätte also in mehr als drei Viertel aller Fälle mit hoher Wahrscheinlichkeit zu unerwünschten Nebenwirkungen geführt.

## Botsis et al. 2010

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Norwegen, USA
3. Bezug zum Thema: Qualitätsindikatoren, SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Elektronische Patientenakten, abgelegt im Klinischen Data Warehouse der Columbia Universität in New York; 3.068 Patienten mit Pankreas-Tumor im Zeitraum 1999 bis 2009.
6. Qualitätsindikatoren: Inkonsistenz, Unrichtigkeit, Unvollständigkeit
7. Schlussfolgerungen der Autoren
   Klinische Register und der Austausch von Gesundheitsdaten können die Datenqualität von elektronischen Patientenakten verbessern. Fortgeschrittene Datenvalidierungswerkzeuge sollten zur Verbesserung der Datenintegrität entwickelt werden.
8. Zusammenfassung der Publikation
   Die Autoren beschreiben verschiedene Datenqualitätsprobleme, auf die sie bei der Überlebenszeitanalyse von Patienten mit Pankreas-Tumor stießen. Aus den elektronischen Patientenakten, die im Klinischen Data Warehouse der Columbia Universität in New York gespeichert sind, wurden anhand des ICD-9-Kodes 3.068 Patienten mit Pankreas-Tumor im Zeitraum 1999 bis 2009 identifiziert. Für diese Patienten wurden zusätzlich Informationen aus verschiedenen Datenelementen der elektronischen Patientenakte (z.B. Pathologie-Befunde, Radiologie-Berichte, Laborwerte) extrahiert. Zur Überprüfung der Datenqualität und Übereinstimmung zwischen verschiedenen Datenelementen wurden folgende Qualitätsindikatoren für einzelne Studienvariablen berechnet.
   - Inkonsistenz: Diskrepanz zwischen verschiedenen oder innerhalb des gleichen Datenelements der elektronischen Patientenakte.
   - Unrichtigkeit: unspezifische, nicht dem Standard folgende, falsche oder unpräzise Information.
   - Unvollständigkeit: fehlende Information.

   Fehlende Informationen bei Schlüsselvariablen führten zum Ausschluss von 1.479 der 3.068 Patienten von der Analyse. Bei einzelnen Studienvariablen erreichte die Unvollständigkeitsrate Werte oberhalb von 50%. Inkonsistenzen wurden sowohl zwischen verschiedenen Datenelementen als auch innerhalb eines Datenelements (z.B. gleichzeitige Angabe

von Diabetes Typ 1 und Typ 2) gefunden. Diese Diskrepanzen ließen sich zum Teil durch redundante Datenerfassung erklären. Unrichtige Daten wurden häufig beobachtet (z.B. Benutzung unspezifischer Diagnose-Kodes, offensichtlich falsche Zeitangaben zur Therapie).

9. Kommentar
Die Integration verschiedener Datenquellen verlangt nach adäquaten Maßnahmen zur Sicherstellung einer guten Datenqualität. Dies betrifft Kohorten und Register in besonderem Maße.

## Bray/Parkin 2009

1. Art des wissenschaftlichen Vorhabens: systematisches Review
2. Herkunft der Autoren: Norwegen
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: Aktualität, Validität, Vergleichbarkeit
7. Schlussfolgerungen der Autoren
Die Vergleichbarkeit von Krebsregisterdaten kann durch einen umfassenden Review des Registrierungsprozesses gefördert werden. Zur Beurteilung der Validität stehen verschiedene Qualitätsindikatoren zur Verfügung. Für die Sicherstellung einer angemessen Aktualität von Krebsregisterdaten gibt es derzeit (2009) keine international verbindlichen Guidelines.
8. Zusammenfassung der Publikation
Dieser Artikel ist der erste Teil eines zweiteiligen Reviews, der praktische Aspekte und Techniken der Datenqualität bei Krebsregistern behandelt. Er fokussiert auf drei Schlüsselaspekte: die Vergleichbarkeit, Validität und Aktualität von Krebsregisterdaten. Vergleichbarkeit ist hier definiert als das Ausmaß, in dem die Prozesse der Kodierung und Klassifizierung sowie die Regeln für die Erfassung und Auswertung spezifischer Kenngrößen international etablierten Standards (z.B. Guidelines) folgen. Für die Vergleichbarkeit spielen vier Aspekte eine wichtige Rolle:
   - das für die Klassifikation und Kodierung von Tumoren benutzte Kodiersystem,
   - die Definition der Inzidenz, d.h. die Definition eines „Falles" und seines Startdatums,
   - die Unterscheidung zwischen Primärfällen und sonstigen Fällen (z.B. Metastasen) sowie
   - die Erfassung von Tumoren bei asymptomatischen Patienten.

   Validität bezeichnet den Anteil der Fälle in einem Register, die eine bestimmte Eigenschaft definitiv besitzen, bezogen auf alle Fälle im Register, für die diese Eigenschaft (korrekt oder fälschlich) dokumentiert ist. Die Validität hängt ab von der Genauigkeit der benutzten Datenquellen und der Erfahrung beim Extrahieren und Kodieren von Informationen

und kann mit verschiedenen Indikatoren gemessen werden (z.B. Rate der DCO-Fälle [„death certificate only“]).
Die Aktualität von Krebsregisterdaten wird durch die Vollständigkeit und Genauigkeiten der Daten beeinflusst. Informationen zu einer Krebserkrankung treffen häufig erst mit einem gewissen Zeitverzug nach Erstdiagnose im Register ein und beeinflussen somit auch die Vollständigkeit der Daten.
9. Kommentar
Die in dieser umfassenden Darstellung vorgestellten Methoden und Verfahren lassen sich teilweise auch für Register nutzen, die nichts mit Krebserkrankungen zu tun haben.

## Brender et al. 2008

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: USA
3. Bezug zum Thema: SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: National Birth Defects Prevention Study (NBDPS) in den USA; daraus 649 Geburten in Texas im Zeitraum 1997–2000.
6. Qualitätsindikatoren: Übereinstimmung
7. Schlussfolgerungen der Autoren
Die Qualität der elterlichen Berufsangaben in den Geburtszertifikaten schwankt sehr und schränkt damit die Verwendbarkeit dieser Information ein.
8. Zusammenfassung der Publikation
In den US-amerikanischen Geburtszertifikaten sollen nach einer Empfehlung des National Center for Health Statistics Angaben über die berufliche Situation der Eltern erfasst werden. Um die Qualität dieser Angaben zu überprüfen, wurden die Angaben aus den Geburtszertifikaten mit Informationen aus Interviews mit den Müttern verglichen (Goldstandard). Zur Beurteilung der Übereinstimmung wurden Sensitivität, Spezifität und die Kappa-Statistik berechnet. Die Übereinstimmung betrug bei den Angaben zur Berufstätigkeit der Mutter 77% und für die Väter 63%. Abhängig von der Berufsgruppe variierte die Sensitivität des Geburtszertifikats zwischen 35% und 100% und die Spezifität zwischen 93% und 100%. Die Kappa-Statistik lag für Mütter höher als für Väter (0,59–0,94 vs. 0,47–0,92 in den verschiedenen Berufsgruppen).
9. Kommentar
Dieser Artikel zeigt den Nutzen von SDV für die Verbesserung der Datenqualität.

## Bronnert et al. 2012

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: USA
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: Aktualität, Definition, Erreichbarkeit, Genauigkeit, Granularität, Präzision, Relevanz, Vollständigkeit, Widerspruchsfreiheit, Zeitnähe
7. Schlussfolgerungen der Autoren
   Mit zunehmenden Datenmengen in der Medizin nimmt auch die Rolle des Health Information Management (HIM) zu. Gutes Data Quality Management (DQM) ist Voraussetzung für hohe Datenqualität und der darauf basierenden Datenanalyse.
8. Zusammenfassung der Publikation
   In dieser Arbeit, die ein Update zu einem bereits bestehenden DQM-Modell darstellt, geben Bronnert et al. Empfehlungen zu einem gutem DQM. DQM ist das Wissen über die Qualität des jeweiligen Datenbestandes sowie die kontinuierliche Verbesserung der Datenqualität in verschiedenen Bereichen. Die Beachtung der Datenqualität beginnt bereits bei der Planung des Einsatzes von Electronic Health Records (EHRs). Es sollten feste Standards und akzeptierte Wertebereiche für die Datenelemente verwendet werden. Bei manueller Datenerfassung sollte das Personal regelmäßig geschult werden. Datenerfassungssysteme sollten regelmäßig kontrolliert werden und vorab definierte Anforderungen erfüllen. Bei der Datenanalyse ist darauf zu achten, dass die Daten im richtigen Kontext verwendet werden und die erforderliche liegende Datenqualität gegeben ist. Bei einem Datawarehouse sollte bekannt sein, wann und wie Datenaktualisierungen bzw. eine Datenpflege durchgeführt werden. Verwendete Versionen von EHRs und anderen Anwendungen, die sich über die Zeit hinweg weiterentwickeln, sollten stets dokumentiert werden. Als wichtige Punkte werden zudem die sichere Verwahrung der Daten und die Ermöglichung eines leichten und effektiven Datenzugangs genannt. Weiterhin präzisieren Bronnert et al. ihr DQM-Modell und bezeichnen Anwendung, Analyse, Erhebungund Verwahrung von Daten als die Domänen der Datenqualität. Als Charakteristiken nennen sie Aktualität, Definition, Erreichbarkeit, Genauigkeit, Granularität, Präzision, Relevanz, Vollständigkeit, Widerspruchsfreiheit und Zeitnähe. Die Begriffe werden in der Arbeit weiterführend definiert.
9. Kommentar
   Die Ausführungen der Autoren sind schlüssig und beleuchten viele Probleme im Umgang mit Datenqualität. Eine ungewöhnliche Strukturierung erschwert leider die Lesbarkeit. Da es sich bei dieser Arbeit unter



anderem um ein Update eines bereits bestehenden DQM-Modells aus dem Jahr 1998 handelt, sind zum besseren Verständnis Kenntnisse der Vorgängerversion notwendig.

## Brower et al. 2006

1. Art des wissenschaftlichen Vorhabens: systematisches Review
2. Herkunft der Autoren: Niederlande
3. Bezug zum Thema: Feedback, Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: Korrektheit, Vollständigkeit
7. Schlussfolgerungen der Autoren
   Die vorhandenen Ansätze zur Verbesserung der Datenqualität von Routinedaten in der hausärztlichen Versorgung bedürfen einer gründlicheren als der bisher (Stand 2006) durchgeführten Untersuchung. Dabei sollte auf eine ausreichend detaillierte Beschreibung der Methoden geachtet werden und sowohl die Vollständigkeit als auch die Korrektheit der Daten bestimmt werden.
8. Zusammenfassung der Publikation
   Die Autoren haben die Frage untersucht, mit welchen Methoden die Datenqualität von Routinedaten in der hausärztlichen Versorgung effektiv verbessert werden kann. Dazu haben sie die Medline-Datenbank mit passenden Begriffen und Schlagworten (MeSH Medical Subject Headings) durchsucht und insgesamt 12 Studien gefunden. Keine einzige Studie untersuchte SDV als Mittel zur Qualitätsverbesserung. Über den Einsatz von Qualitätsindikatoren wurde in 10 der 12 Studien berichtet. Die häufigste Methode zur Qualitätsverbesserung war ein individualisiertes Feedback. Fast alle in den Studien betrachteten Ansätze zur Verbesserung der Datenqualität waren in irgendeiner Weise mehr oder weniger wirksam.
9. Kommentar
   Diese Arbeit zeigt, dass – Stand 2006 – noch ein erheblicher Forschungsbedarf im Bereich der Datenqualitätsverbesserung bei der hausärztlichen Routinedokumentation besteht.

## Chiba et al. 2012

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Japan
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Routinedaten zweier ländlicher Krankenhäuser in Kenia
6. Qualitätsindikatoren: Genauigkeit, Relevanz, Vollständigkeit

7. Schlussfolgerungen der Autoren
   Die Datenqualität zweier Geburtenregister ländlicher Krankenhäuser erwies sich bei einer qualitativen und quantitativen Überprüfung als nicht zufriedenstellend. Die Mängel sind auf organisatorische und technische Faktoren sowie Fehlverhalten der Hebammen zurückzuführen. Politisches Engagement ist nötig, um gegen diese Probleme weiter vorzugehen.
8. Zusammenfassung der Publikation
   Die Prävalenz der Müttersterblichkeit ist in Entwicklungsländern wie Kenia erschreckend hoch. Qualitativ hochwertige Daten zu tödlichen Komplikationen während der Geburt können helfen, deren Relevanz zu erkennen, gegen sie vorzugehen und insgesamt die Versorgung der Mütter zu verbessern. Solche Daten gibt es jedoch in Kenia nicht. Für eine retrospektive vergleichende Studie wurden im Jahr 2006 Daten der Geburtenregister zweier ländlicher Krankenhäuser in den Gebieten Siaya und Bondo untersucht. Der Fokus der Untersuchung lag hierbei auf den Datenqualitätskriterien Relevanz, Vollständigkeit und Genauigkeit. Die Registerdaten wurden auf einen Computer übertragen und dabei einer Prüfung auf Korrektheit unterzogen. Die Gesamtzahl der Einträge wurde zudem mit denen der monatlichen Leistungsstatistiken verglichen. Von den 30 untersuchten Datenfeldern waren in Siaya fünf und in Bondo 23 Kategorien zu mehr als 20% unvollständig. Im Median waren die Daten in Siaya zu 2,75% sowie in Bondo zu 33,85% unvollständig. Zudem wurden generell zu wenige Kategorien in die im A3-Format vorliegenden Register aufgenommen. Nur 66,2% in Bondo bzw. 97,3% der Daten in Siaya waren im Median korrekt. Die Geburtenregister in Siaya und Bondo können zusammenfassend als für die Forschung unangemessen und irrelevant betrachtet werden. Es wurde versucht, die Gründe der Datenmängel zu ermitteln. Hierbei wären grundsätzlich organisatorische Probleme zu nennen, die beispielsweise fehlende einheitliche Werkzeuge zur Datenaufzeichnung bzw. von den Müttern persönlich zu tragenden Kosten der Datenerfassung umfassen. Ein weiterer Faktor ist das fehlerhafte Verhalten der Hebammen bei der Datenerfassung.
9. Bewertung der Publikation
   Mängel in der Qualität von Routinedaten eines Entwicklungslandes wie Kenia wurden über die Überprüfung verschiedener Qualitätskriterien veranschaulicht. Hierbei zeigen sich ein deutlicher Unterschied zu westlichen Standards und die Notwendigkeit weiterer politischer Hilfestellungen.

## Choquet et al. 2010

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Frankreich
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar

5. Datenbestand: Daten des Georges Pompidou European Hospital in Paris/Frankreich
6. Qualitätsindikatoren: nicht anwendbar
7. Schlussfolgerungen der Autoren
   Die vorgestellte Methodologie erlaubt eine Verallgemeinerung über den untersuchten Anwendungsbereich hinaus.
8. Zusammenfassung der Publikation
   Die Autoren beschreiben einen Ansatz zur Untersuchung und Beeinflussung der Datenqualität bei der Integration von elektronischen Patientenakten in ein klinisches Data Warehouse. Startpunkt ist ein „Dreieck der Informationsqualität“ mit den 3 Ecken „Objekte“ (Daten über Objekte der realen Welt), „Konzepte“ (Informationsmodelle) und „Terms“ (Terminologie, Thesaurus, Klassifikation oder Vokabular). Für jede „Ecke“ gibt es Methoden und Scores, die jeweils die Datenqualität messen. Diese Scores werden in einem globalen Score zur Beschreibung der Datenqualität einer Informationsquelle zusammengefasst. Eine praktische Anwendung dieses Ansatzes erfolgte an den Daten des Georges Pompidou European Hospital in Paris/Frankreich im Rahmen des EU FP7 Projektes DebugIT. Dieses Projekt soll eine technische und semantische Plattform aufbauen, um den Datenaustausch zwischen europäischen Krankenhäusern über Infektionskrankheiten und antimikrobielle Resistenzen zu unterstützen.
9. Kommentar
   Der „Dreieck der Informationsqualität“ scheint in der Anwendung recht aufwendig zu sein. Seine Eignung für Register und Kohortenstudien müsste erst noch gezeigt werden.

## Couchoud et al. 2013

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Frankreich
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: Aktualität, Brauchbarkeit, Duplizität, Empfindlichkeit, Genauigkeit, Validität, Vergleichbarkeit, Vollständigkeit, Vollzähligkeit
7. Schlussfolgerungen der Autoren
   Für die Untersuchung und Verbesserung der Datenqualität in Dialyse-Registern können Methoden und Indikatoren benutzt werden, die bei longitudinalen Studien zum Einsatz kommen.
8. Zusammenfassung der Publikation
   Die Autoren beschreiben verschiedene Aspekte der Datenqualität in Bezug auf Dialyse-Register. Folgende Dimensionen der Datenqualität (Qualitätsindikatoren) sind für solche Register besonders bedeutsam.

- **Vollzähligkeit**: alle Dialyse-Patienten, die die Einschlussbedingungen des Registers erfüllen, und alle ihre Ereignisse sind im Register gespeichert
- **Vollständigkeit** auf Merkmals-Ebene: die Rate fehlender Werte pro Merkmal sollte möglichst gering sein
- **Vermeidung** von Duplikaten: jeder Patient/jede Patientin sollte nur einmal im Register gespeichert sein
- **Validität** der Merkmale: Anteil der Fälle im Register mit einer dokumentierten Eigenschaft, die diese Eigenschaft tatsächlich aufweisen
- **Exaktheit** der Datenelemente: die Datenelemente bilden genau die Merkmale ab, die gemessen werden sollen
- **Vergleichbarkeit** der Daten mit anderen Registern
- **Brauchbarkeit** der Merkmale für aktuelle und zukünftige Auswertungen
- **Empfindlichkeit** der Merkmale: Fähigkeit eines Merkmales, Änderungen widerzuspiegeln
- **Aktualität** der Daten
- **Exaktheit** der Datenauswertungen und **Interpretation** der Ergebnisse

Zur Schätzung einiger dieser Qualitätsindikatoren gibt es als Goldstandard Empfehlungen der North American Association of Central Cancer Registries.

- Zur **Vollzähligkeit**: Vergleich mit anderen Registern; Re-Identifizierung der Fälle aus einer vom Register unabhängigen Quelle
- Zur **Vollständigkeit** auf Merkmals-Ebene: Berechnung des Anteils der Fälle mit fehlenden Angaben bezogen auf alle registrierten Fälle
- Zur **Vermeidung** von Duplikaten: Regeln zur Identifikation von Individuen; regelmäßige Suche nach Duplikaten anhand personenbezogener Daten (z.B. Name, Geburtsdatum), sofern datenschutzrechtlich erlaubt
- Zur **Validität** der Merkmale und der **Exaktheit** der Datenelemente: Doppelerfassung von Merkmalen durch verschiedene Personen und Prüfung der Übereinstimmungsrate; Plausibilitätsprüfungen
- Zur **Vergleichbarkeit** der Daten: Formale Beschreibung aller benutzten Definitionen und Regeln
- Zur **Brauchbarkeit** der Merkmale: Für jedes Merkmal die Berechnung der Anzahl wissenschaftlicher Publikationen des Registers, die dieses Merkmal in irgendeiner Form nutzen
- Zur **Aktualität** der Daten: Berechnung der Zeitdifferenz zwischen klinischem Ereignis und seiner Erfassung im Register; Berechnung der Zeitdifferenz zwischen Erfassung im Register und der Bereitstellung für wissenschaftliche Auswertungen

9. Kommentar
   Die vorgestellten Qualitätsindikatoren lassen sich auch für andere Arten von Registern sowie Kohortenstudien einsetzen.




## De 2011

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Indien
3. Bezug zum Thema: SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: nicht anwendbar
7. Schlussfolgerungen der Autoren
   Es ist an der Zeit, die 100%-SDV aus der Zeit der papiergestützten CRFs durch eine Kombination neuer Methoden zu ersetzen. Dies führt zu einer Einsparung von Ressourcen, die an anderer Stelle bei der Studienplanung und -durchführung eingesetzt werden können.
8. Zusammenfassung der Publikation
   Der Autor beschreibt Möglichkeiten, bei klinischen Studien auf eine vollständige SDV zu verzichten und trotzdem die regulatorischen Anforderungen an die Qualität klinischer Studien zu erfüllen, wie sie etwa in der ICH Guideline E6 Good Clinical Practice niedergelegt sind. Folgende Methoden können einzeln oder in Kombination genutzt werden:
   - „Zielgerichtete“ SDV: Mittels statistischer Verfahren werden die zu prüfenden Patienten und Daten-Elemente ausgewählt.
   - Risikobasierte SDV: Die SDV wird auf besonders risikobehaftete Daten-Elemente konzentriert, d.h. auf Daten-Elemente, bei denen Fehler eine große Auswirkung auf die Qualität der Studiendaten oder -ergebnisse aufweisen
   - Adaptive SDV: Ein initialer Monitoring-Plan wird dynamisch an die Ergebnisse der durchgeführten SDV angepasst (z.B. Konzentration auf Studienzentren mit besonders schlechter Datenqualität)
   - SDV „aus der Ferne“: SDV-Aktivitäten, die üblicherweise vor Ort in den Studienzentren durchgeführt werden, werden in die Studienzentrale verlagert
   - Hybride SDV: Eine Mischung aus „zielgerichteter“ SDV (Punkt 1) und SDV „aus der Ferne“ (Punkt 4)
9. Kommentar
   Die hier beschriebene adaptive SDV entspricht dem Konzept, das die Leitlinie vorschlägt.

## Duda et al. 2012

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: USA
3. Bezug zum Thema: SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Caribbean, Central and South America Network for HIV Epidemiology (CCASAnet)

6. Qualitätsindikatoren: Korrektheit, Vollständigkeit
7. Schlussfolgerungen der Autoren
   In Kohortenstudien sollten Audits mit SDV genutzt werden, um die Datenqualität zu prüfen und zu verbessern und den Prozess der Datensammlung zu optimieren.
8. Zusammenfassung der Publikation
   Die Autoren untersuchten mittels SDV bei Vor-Ort-Audits zwischen April 2007 und März 2008 die Korrektheit und Vollständigkeit von Studiendaten, die aus der Routinedokumentation der beteiligten Kliniken gewonnen wurden. Die Kliniken sind Teil eines internationalen HIV-Forschungsnetzwerkes und nutzen unterschiedliche Wege zur Datenübermittlung an das Netzwerk. Jeweils 30 zufällig ausgewählte Datensätze aus dem Datenbestand des Netzwerkes wurden mit den Originaldaten in der untersuchten Klinik verglichen. Das Vergleichsergebnis wurde pro Daten-Element (demographische Angaben, HIV-relevante medizinische Informationen) folgendermaßen klassifiziert:
   - Korrekt: Wert in der Studiendatenbank ist mit Originalwert identisch.
   - Geringfügige Fehler („minor error“): Fehler, der kein „major error“ ist.
   - Bedeutsamer Fehler („major error“): Klinisch bedeutsamer Unterschied zwischen Wert in der Studiendatenbank und Originalwert.
   - Fehlender Wert in der Studiendatenbank: Vorhandener Originalwert wurde nicht in die Studiendatenbank übertragen.
   - Quellenloser Wert: Zu dem Wert in der Studienbank wurde kein Originalwert gefunden.

   Aus den Einzelergebnissen wurden Fehlerraten pro Datenelement berechnet.
   5 der 7 untersuchten Kliniken wiesen Fehlerraten > 10% bei wichtigen Studienvariablen, Labordaten, Gewichtsmessungen und antiretroviralen Medikamenten auf. In allen untersuchten Kliniken gab es signifikante Abweichungen bei dem Beginn- oder Endedatum von Medikamenten.
9. Kommentar
   Die Untersuchung zeigt, dass die SDV bei der Nutzung heterogener Datenquelle eine wichtige Rolle bei der Beurteilung und Verbesserung der Datenqualität spielt.

### Dyck et al. 2007

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: USA
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Studiendaten zu in Altersheimen untergebrachten und unter Delirium leidenden Patienten.
6. Qualitätsindikatoren: Validität, Vollständigkeit, Zuverlässigkeit.

7. Schlussfolgerungen der Autoren
   Die Datenqualität beeinflusst direkt die Zuverlässigkeit und Validität einer Studie. Um diese zu gewährleisten, wurde für die vorliegende Kohortenstudie ein aufwendiges Programm zur Qualitätssicherung der Daten erarbeitet.
8. Zusammenfassung der Publikation
   Die Autoren dieser Arbeit erläutern methodische Anstrengungen, eine hohe Datenqualität im Kontext einer Kohortenstudie mit an Delirium leidenden, in Altersheimen untergebrachten Patienten zu gewährleisten. Die Studie fand in ländlicher Umgebung im Südwesten der USA im Bundesstaat Iowa statt. Die vorgestellte Strategie des Qualitätsmanagements richtete sich in erster Linie nach einer Vorlage von Donabedian et al. Hierbei wurde das Qualitätsmanagement in zwei Teilbereiche, die Konzeption der zugrunde liegenden Struktur sowie die Verwaltung der Prozesse unterteilt. Als strukturelle Komponenten wurden der Studienantrag, die Studienzentren, die Instrumente, Laboreinrichtungen als auch das Personal eingeordnet. Zu jedem Teilbereich wurden dabei aktuelle standardisierte Vorlagen verwendet. Patientendaten wurden durch die zufällige Auswahl der teilnehmenden Pflegeheime gefolgt von der zufälligen Auswahl der dort ansässigen Patienten ermittelt. Auch die Untersuchungsinstrumente richteten sich nach Standards, und die Durchführung der Untersuchungen erfolgte jeweils zu denselben Tageszeiten. Forschungsassistenten waren angewiesen, zeitnah die Daten in Protokolle aufzunehmen. Zudem verfügte die Studie über einen Datenqualitätskoordinator, der die Arbeit der Forschungsassistenten koordinierte und kontrollierte.
   Für das Datenqualitätsprogramm wurden Konzepte zu den folgenden drei Prozeduren ausgearbeitet und umgesetzt: Kompetenzeinschätzung, Bewertung der Reliabilität sowie Berichte zur Datenvollständigkeit. Zur Kompetenzeinschätzung wurden die Forschungsassistenten durch den Datenqualitätskoordinator bezüglich der notwendigen Datenelemente, beispielsweise dem Ausmaß des Deliriums des jeweiligen Patienten, regelmäßig geschult und mit einem schriftlichen Test geprüft. Weitere Maßnahmen hierzu umfassten den wöchentlichen Kontakt mit dem Forschungsvorstand.
   Zur regelmäßigen Prüfung der Reliabilität zwischen den Forschungsassistenten wurde deren Übereinstimmung bei der Bewertung kognitiver Tests verwendet. Lag diese unter 90%, wurde eine Nachschulung angeordnet. Die Qualitätskontrolle zu Vollständigkeit und Korrektheit erfolgte im Forschungsbüro des Studienteams. Werte fehlten meist dann, wenn die Patienten nicht in der Lage waren oder sich weigerten, spezifische Fragen zu beantworten. Die Dateneingabe in die Datenbank erfolgte ausschließlich durch den Forschungsvorstand.
9. Bewertung der Publikation
   Sehr praxisorientiert beschreibt die Arbeit das genaue Vorgehen zur Gewährleistung einer hohen Datenqualität in der vorgestellten Kohorten-

studie. Dabei erscheint besonders die Schulung und Kontrolle des Forschungspersonals zwar als aufwendig, jedoch effektiv und gut durchdacht und auf andere Projekte übertragbar.

### França et al. 2008

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Brasilien
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: brasilianische Todesursachenstatistik aus den Jahren 2002–2004
6. Qualitätsindikatoren: Generalisierbarkeit, Korrektheit, politische Relevanz, Zuverlässigkeit
7. Schlussfolgerungen der Autoren
   Bei Auswertung der Todesursachenstatistik in Brasilien müssen Mängel bei deren Vollständigkeit sowie hinsichtlich der Falschkodierung von Todesursachen beachtet werden. Vor allem in den nördlichen, sozioökonomisch schwachen Regionen sind Maßnahmen notwendig, um die Datenqualität verbessern.
8. Zusammenfassung der Publikation
   Eine der größten Einschränkungen bei der Analyse von Daten der Todesursachenstatistik in Entwicklungsländern ist die Qualität der Daten. Trotz gesetzlicher Maßnahmen stellt sich der betreffende Datenbestand in Brasilien als fragmentiert und inkonsistent zwischen den Regionen dar. Aus diesem Grund wurde eine umfangreiche Bewertungsstruktur geschaffen, um Unterschiede in der Qualität der Daten zwischen den brasilianischen Regionen zu identifizieren. Die zu untersuchenden Daten aus den Jahren 2002 bis 2004 wurden aus dem 1975 gegründeten Mortality Information System bezogen. Die verwendete Bewertungsstruktur, formalisiert in einer vorhergenden Arbeit der Co-Autoren [Rao et al. 2005], untersucht vier Hauptattribute bestehend aus neun zugeordneten Kriterien:
   - Generalisierbarkeit (mit den Kriterien Abdeckung und Vollständigkeit),
   - Korrektheit (Verwendung von unspezifischen Kodes, Inhaltskorrektheit, unrichtige oder unwahrscheinliche Alters- oder Geschlechtsverteilungen),
   - Zuverlässigkeit (generelles Niveau der Sterblichkeit und Konsistenz von ursachenspezifischen Sterblichkeitseigenschaften über die Zeit) sowie
   - politische Relevanz (Aktualität und geographische Nicht-Übereinstimmung).

Die Untersuchung der Generalisierbarkeit befasste sich mit der Frage, ob das Mortality Information System tatsächlich die Daten aller Regionen vollständig abdeckt. Hierfür wurde anhand von Volkszählungsdaten geprüft, ob die Daten der Todesursachenstatistik unterrepräsentiert vorliegen. Auf nationaler Ebene beträgt die Vollständigkeit der registrierten Todesfälle ca. 90%. Unterschiede zwischen den Regionen waren jedoch vorhanden und sind auf die ökonomischen Unterschiede zurückzuführen. Zur Prüfung der Korrektheit wurden die gegebenen Diagnosekodes der Jahre 2002 bis 2004 manuell auf Stimmigkeit geprüft. Hierbei wurden je nach Region im Mittel 13,1% der ermittelten Diagnosekodes als unstimmig eingeordnet. Anhand der beobachteten und der nach Alter und Geschlecht vorhergesagten Sterblichkeit bei Tod durch Fremdeinwirkung wurde die Zuverlässigkeit geprüft. Aus dieser Untersuchung ließen sich jedoch keine Probleme der Datenqualität ableiten. Bei der Untersuchung der politischen Relevanz zeigte sich, dass die Verfügbarkeit der Daten sowohl zeitlich (ca. 18 Monate Verzug) als auch angemessen für subnationale Analysen gegeben ist. Grundsätzlich ließ sich beobachten, dass die Datenqualität der Daten des Mortality Information Systems in den südlichen, südöstlichen sowie zentralwestlichen Regionen befriedigend ist. Die Datenqualität der nördlichen bzw. nordöstlichen Regionen zeigt jedoch Mängel.

9. Kommentar
   Anhand einer umfangreichen Bewertungsstruktur wurden anschaulich verschiedene Aspekte zur Datenqualität der brasilianischen Todesursachenstatistik untersucht. Hierbei zeigt sich eine starke Korrelation zwischen Datenqualität und sozioökonomischem Niveau der einzelnen Regionen.

## Kahn et al. 2012

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: USA
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: angemessenes Datenvolumen, Glaubwürdigkeit, Korrektheit, Objektivität, Zeitnähe
7. Schlussfolgerungen der Autoren
   Eine systematische Vorgehensweise bei der Bewertung von Datenqualität ist essentiell, um das Potenzial elektronischer Patientenakten für die Forschung ausschöpfen zu können.
8. Zusammenfassung der Publikation
   Daten aus elektronischen Patientenakten werden zunehmend für wissenschaftliche Studien genutzt. Die Autoren schlagen ein Konzeptmo-

dell vor, um die Datenqualität solcher Akten, insbesondere wenn sie als Quelle für multizentrische Studien dienen, zu prüfen. Dieses Konzeptmodell ist eine Vereinfachung eines früher von anderen Autoren vorgeschlagenen Modells und umfasst zwei Kategorien:

- Intrinsisch: Merkmale von Datenqualität, die den Daten selbst zuzuordnen sind (Qualitätsindikatoren Korrektheit, Objektivität, Glaubwürdigkeit).
- Konzeptuell: Merkmale von Datenqualität, die im Kontext der geplanten Datennutzung relevant sind (Qualitätsindikatoren Rechtzeitigkeit, angemessenes Datenvolumen).

Ergänzt wird dieses Modell durch ein Prozess-Modell für die Planung und Durchführung von unizentrischen und multizentrischen Datenqualitätsprüfungen. Dazu gehören eine Priorisierung von Variablen und Datenqualitätsdimensionen, Standardverfahren für die Prüfung der Datenqualität und iterative Prüfungen der Datenqualität innerhalb von Studienzentren und zentrumsübergreifend. Die Aktivitäten sollten auf die Datenbereiche konzentriert werden, die besonders anfällig für Qualitätsprobleme sind, und umfassend dokumentiert werden.

9. Kommentar
Die Autoren fassen bestehende Erkenntnisse in einem neuen Modell zusammen.

## Krzych et al. 2011

1. Art des wissenschaftlichen Vorhabens: Analyse
2. Herkunft der Autoren: Polen, Vereinigtes Königreich
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: multi-zentrische Studie „Arterial Revascularisation Trial (ART)“, Daten aus verschiedenen Ländern
6. Qualitätsindikatoren: Aktualität, Rekrutierungsrate, Vollständigkeit
7. Schlussfolgerungen der Autoren
Die Studie zeigte eine vergleichbare Datenqualität in entwickelten und Entwicklungsländern. Wichtig sind ein angemessenes Training der Studienzentren und ein zentrales Management der Datenqualität.
8. Zusammenfassung der Publikation
Die Autoren untersuchten die Datenqualität einer multi-zentrischen randomisierten Studie anhand der Qualitätsindikatoren Rekrutierungsrate, Aktualität und Vollständigkeit in verschiedenen Studienzentren in Ländern mit unterschiedlichem sozioökonomischen Status. Die Analyse basierte auf den CRFs, die zum 6-Wochen-Follow-Up der Studie ausgefüllt werden mussten. Die ausgefüllten CRFs wurden danach unterschieden, ob Rückfragen („Queries“) notwendig waren. Die Aktualität der Daten wurde bestimmt durch die Zeit zwischen Durchführung des Fol-

low-Ups und Erhalt der ausgefüllten CRFs in der Studienzentrale. Berechnet wurden die prozentuale Rate von Queries und die Anzahl Tage zwischen Follow-Up und Erhalt der korrekten Daten. Beim sozioökonomischen Status wurde gemäß einer Definition der Weltbank unterschieden zwischen entwickelten Ländern (Italien, Großbritannien, Österreich, Australien; insgesamt 18 Studienzentren) und Entwicklungsländern (Polen, Brasilien, Indien; insgesamt 10 Studienzentren). Die Rekrutierungsrate unterschied sich nicht signifikant zwischen Entwicklungsländern (4,1 Personen/Monat) und entwickelten Ländern (3,7 Personen/Monat). In den Entwicklungsländern dauerte es im Median 37 Tage, bis die Daten des Follow-Ups in der Studienzentrale eintrafen; in den entwickelten Ländern waren es im Median 11 Tage. Der Median der Query-Rate lag in den Entwicklungsländern bei 19% gegenüber 23% in den entwickelten Ländern.

9. Kommentar
   Der Einfluß sozioökonomischer Faktoren auf die Datenqualität wird sehr selten untersucht. Diese Untersuchung ist bedeutsam für multi-nationale Kohorten und Register.

## Larsen et al. 2009

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Norwegen
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Norwegisches Krebsregister
6. Qualitätsindikatoren: Aktualität, Korrektheit, Vergleichbarkeit, Vollständigkeit
7. Schlussfolgerungen der Autoren
   Das norwegische Krebsregister enthält Daten, die international vergleichbar und angemessen vollständig und richtig sind. Damit ist es gerechtfertigt, jährliche Inzidenzen ein Jahr nach der Diagnose zu berichten.
8. Zusammenfassung der Publikation
   Ziel der Untersuchung war die Bewertung der Datenqualität im Norwegischen Krebsregister. Dazu wurden die Qualitätsindikatoren Vergleichbarkeit, Vollständigkeit, Korrektheit und Aktualität für den Zeitraum 1953–2005 mit besonderem Augenmerk auf dem Zeitraum 2001–2005 analysiert. Das norwegische Krebsregister folgt bei der Kodierung und Klassifikation internationalen Standards und erlaubt somit einen Vergleich seiner Daten mit anderen Krebsregistern. Die Vollständigkeit lag im Zeitraum 2001–2005 insgesamt bei 98,8%. Hämatologische bösartige Tumore (ICD-10-Kodes C90–C95) und Tumore des zentralen Nervensystems (ICD-10-Kodes C70–C72) erreichten eine deutlich schlechtere Vollständigkeit.

Im Zeitraum 2001–2005 wurden 93,8% der Fälle morphologisch verifiziert. Der Median der Zeitdifferenz zwischen Diagnosedatum und Datum der Registrierung im Krebsregister reduzierte sich von mehr als 525 Tagen im Jahre 2001 auf 261 Tage im Jahre 2005.

9. Kommentar
   Die international etablierten Standards für Krebsregister scheinen sich in Punkto Datenqualität zu bewähren.

### Loane et al. 2011

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Vereinigtes Königreich
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Register kongenitaler Anomalien, zusammengeschlossen im Netzwerk European Surveillance of Congenital Anomalies (EUROCAT).
6. Qualitätsindikatoren: Korrektheit, Verfügbarkeit, Vollständigkeit, Vollzähligkeit, Zeitnähe
7. Schlussfolgerungen der Autoren
   Die von EUROCAT benutzten Datenqualitätsindikatoren erlauben einen Vergleich der teilnehmenden Register und geben Hinweise auf Verbesserungsmöglichkeiten bei der Datenerfassung.
8. Zusammenfassung der Publikation
   Das europäische Netzwerk von Registern zu kongenitalen Anomalien (EUROCAT = European Surveillance of Congenital Anomalies) sammelt Daten zu Lebendgeburten, fetalem Tod ab der 20. Schwangerschaftswoche und Schwangerschaftsabbrüchen wegen fetaler Anomalien. Zur Sicherstellung einer hohen Datenqualität werden 30 Qualitätsindikatoren benutzt, die sich auf folgende Schlüsselelemente der Datenqualität beziehen:
   - Korrektheit der Diagnose (12 Qualitätsindikatoren),
   - Verfügbarkeit von Bevölkerungszahlen für die Berechnung von Raten (2 Qualitätsindikatoren),
   - Vollzähligkeit (10 Qualitätsindikatoren),
   - Vollständigkeit der Informationen in den EUROCAT-Variablen (5 Qualitätsindikatoren) sowie
   - Zeitnähe der Datenübermittlung (1 Qualitätsindikator).

   Die Werte der Qualitätsindikatoren werden für jedes beteiligte Register und für den gesamten Datenbestand berechnet und auf der Homepage von EUROCAT präsentiert.
9. Kommentar
   Die Datenqualitätsindikatoren sind so speziell auf kongenitale Anomalien ausgerichtet, dass sie nicht für andere Register oder Kohorten verallgemeinert werden können.

## Macefield et al. 2013

1. Art des wissenschaftlichen Vorhabens: systematisches Review
2. Herkunft der Autoren: Vereinigtes Königreich
3. Bezug zum Thema: SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: nicht anwendbar
7. Schlussfolgerungen der Autoren
   Es gibt eine große Vielfalt unterschiedlicher Verfahrensweisen für das Vor-Ort-Monitoring.
8. Zusammenfassung der Publikation
   Die Autoren untersuchten in der Literatur beschriebene Verfahren zur SDV mittels Vor-Ort-Monitoring bei randomisierten Studien des Gesundheitswesens. Die mittels systematischer Literaturrecherche gefundenen Quellen wurden aufgeteilt in die Kategorien:
   - Berichte von Forschungsgruppen und -organisationen
   - Berichte von einzelnen randomisierten Studien
   - Randomisierte Studien, in denen Varianten des Vor-Ort-Monitorings getestet wurden
   - Kostensimulationen
   - Umfragen bei Studienmitarbeitern und Monitoren

   Informationen aus Quellen der Kategorien 1 bis 3 wurden zusammengefasst hinsichtlich der Häufigkeit und des Umfangs der Monitorbesuche, der Größe und Zusammensetzung des Monitoring-Teams, der Tätigkeiten während der Monitorbesuche sowie des Berichtswesens. Vorteile und Nachteile des Vor-Ort-Monitorings wurden für alle betrachteten Quellen ausgewertet. 57 relevante Artikel wurden identifiziert (Kategorie 1: 21 Artikel, Kategorie 2: 30 Artikel, Kategorie 3: 1 Artikel, Kategorie 4: 2 Artikel, Kategorie 5: 3 Artikel). Die Besuchshäufigkeit des Vor-Ort-Monitorings variierte zwischen 6–8 Wochen und 3 Jahren, die Anzahl der Besuche zwischen 1 und 8. In den meisten Fällen wurde ein SDV durchgeführt und die Einverständniserklärungen geprüft. Als Vorteile des Vor-Ort-Monitorings wurden verbesserte Rekrutierungsraten und eine bessere Einhaltung des Studienprotokolls berichtet. Als Nachteile erwiesen sich die direkten Kosten und die benötigte Arbeitszeit der Monitore.
9. Kommentar
   Eine systematische Evaluation des Vor-Ort-Monitorings und der dabei durchgeführten SDV fehlt bisher. Der vorliegende Artikel ist ein wichtiger Ansatz, das publizierte Wissen zu dieser Thematik zusammenzufassen.

## Maruszewski et al. 2005

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: international
3. Bezug zum Thema: Qualitätsindikatoren, SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Datenbestand der European Congenital Heart Defects Database (ECHSA) aus dem Jahr 2003.
6. Qualitätsindikatoren: Genauigkeit, Vollständigkeit
7. Schlussfolgerungen der Autoren
   Bei der SDV in verschiedenen europäischen Zentren zeigten sich keine statistisch signifikanten Unterschiede zwischen nicht-verifizierten und verifizierten Daten. Allerdings leidet die Studie unter der willkürlichen Auswahl der Datenzentren, bei denen es sich um Standorte handelte, die bereits als qualitativ hochwertige Datenlieferanten bekannt waren.
8. Zusammenfassung der Publikation
   Die European Association for Cardio-Thoracic Surgery (EACTS) trägt derzeit die finanzielle Verantwortung für eine seit den 1990er Jahren bestehende Datenbank, in der europaweit Angaben zu Operationen bei angeborenen Herzfehlern gesammelt werden. Das Management der Datenbank entschied sich für einen Datenverifikationsversuch mittels Source Data Verification (SDV).
   In fünf europäischen Zentren, die sich freiwillig gemeldet hatten, wurden im Zeitraum vom 01.06.2004 bis 31.07.2004 die Daten von insgesamt 32% der Patienten und 34% der Prozeduren aus dem Jahre 2003 von zwei Mitarbeitern der Datenbank geprüft. Die Prüfung umfasste folgende Datenelemente: 30-Tage-Mortalität, Länge des Aufenthaltes im Zentrum, Geburtsdatum, Datum der Aufnahme, Datum der Operation, Datum der Entlassung/des Todes, Körpergewicht, Fallkategorie, Zeitdauer des kardiopulmonalen Bypasses, Zeitdauer der Aortaklemme und Zeitdauer des Kreislaufstillstands.
   Verglichen wurden die Mittelwerte der einzelnen Datenelemente vor und nach der Datenverifikation mittels Welch’s t-Test. Statistische Signifikanz wurde bei einem p-Wert < 0,05 angenommen. Es wurden keine statistisch signifikanten Unterschiede zwischen den nicht-verifizierten und den verifizierten Daten festgestellt.
9. Bewertung der Publikation
   Die Ergebnisse der Studie illustrieren die Notwendigkeit, bei Datenqualitätsanalysen möglichst jede Befangenheit zu vermeiden. Zwar erscheinen die Ergebnisse sehr positiv; durch die gezielte Auswahl der geprüften Standorte kommen jedoch Zweifel auf, ob die Ergebnisse verallgemeinerbar sind. Eine zufällige Auswahl des Standortes oder eine SDV über den Datenbestand aller assoziierten Zentren aus dem Jahr 2003 hätte diese Problematik beseitigt.

## McKenzie et al. 2005

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Australien
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Patientendaten des Queensland Trauma Registry (QTR) aus dem Jahr 1998 sowie die dazugehörigen Ursprungsdaten mit passenden Einschlusskriterien eines Krankenhauses aus Brisbane.
6. Qualitätsindikatoren: Konkordanz
7. Schlussfolgerungen der Autoren
   Obwohl eine Konkordanz von 95% zwischen Krankenhausdaten und Registerdaten relativ gut erscheint, legt die Studie offen, dass der Prozess der Datenaufnahme in die Register verbessert werden kann. Regelmäßige Kontrollen zur Qualität und Vollständigkeit sind wichtig, wenn Registerdaten für die medizinische Forschung verwendet werden sollen.
8. Zusammenfassung der Publikation
   Im Jahr 2003 wurde an einem Krankenhaus in Brisbane eine Studie zu Sterberate und Todesursache bis zu vier Jahre nach Krankenhausaufnahme bei Patienten mit schweren Verletzungen durchgeführt. Unter Erfüllung bestimmter Einschlusskriterien wurden die Daten der betroffenen Patienten in das QTR übertragen. Die Einschlusskriterien waren dann erfüllt, wenn Patienten im Jahr 1998 für mindestens 24 Stunden aufgrund einer Verletzung stationär behandelt wurden. Anhand dieser Studie wurden Konkordanz sowie Typ und Quelle von Datenqualitätsproblemen zwischen den Daten des QTR und den ursprünglich erhobenen Krankenhausdaten bestimmt.
   Aus den aus 1998 verfügbaren QTR-Daten wurden hierzu demografische Variablen zu Namen, Geschlecht, Geburtsdatum sowie, falls verfügbar, Sterbedatum exportiert und die dazu passenden, im Krankenhaus verfügbaren Ursprungsdaten bestimmt. Hierbei wurden noch einmal anhand der Krankenhausdaten die Einschlusskriterien, speziell Aufnahme- und Entlassungsdatum, geprüft. Zudem wurden gegenläufig die Krankenhausdaten retrospektiv auf Patienten mit passenden Einschlusskriterien geprüft, die durch das Matching mit den QTR-Daten nicht identifiziert werden konnten.
   Folgende vier Hauptklassen von Problemen wurden identifiziert:
   - Scheitern der Wiedererkennung relevanter Patienten anhand der QTR-Daten,
   - fälschliche Aufnahme von irrelevanten Patientendaten in das QTR,
   - unvollständige bzw. ungenaue Krankenhausdaten sowie
   - unvollständige bzw. ungenaue Datenaufnahme des QTR

   Informationen zu 1.749 Patienten waren in den QTR-Daten vorhanden. Hiervon erfüllten 71 Patienten (ca. 4%) nicht die Einschlusskriterien.

Zwölf Patienten wurden vom QTR fälschlicherweise nicht registriert (ca. 1%). Die Konkordanz zwischen den Daten des Krankenhauses sowie des QTR liegt bei ca. 95%. In 52% der Fälle beruhten die Unstimmigkeiten auf Fehlern des QTR, in 22% der Fälle auf Fehlern des Krankenhauses und in 26% der Fälle war die Fehlerquelle unsicher.

9. Kommentar
   Die Arbeit zeigt, dass trotz festgelegter Einschlusskriterien Unterschiede zwischen Register- und Krankenhausdaten auftreten können. Gerade die Interpretation bzw. Kommunikation der Einschlusskriterien mag in dieser Arbeit für die Unterschiede gesorgt haben und zeigt folglich die Notwendigkeit einer strikten Definition.

### Messenger et al. 2012

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: USA
3. Bezug zum Thema: Feedback, Qualitätsindikatoren, SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: National Cardiovascular Data Registry (NCDR) in den USA
6. Qualitätsindikatoren: Korrektheit, Vollständigkeit, Widerspruchsfreiheit
7. Schlussfolgerungen der Autoren
   Die Ergebnisse des Audits von 2010 zeigen eine hohe Übereinstimmung der Registerdaten mit den Originaldaten.
8. Zusammenfassung der Publikation
   Das US-amerikanische Register für kardiovaskuläre Daten (National Cardiovascular Data Registry [NCDR]) hat ein Datenqualitätsprogramm entwickelt, um die Vollständigkeit, Widerspruchsfreiheit und Richtigkeit der Daten zu gewährleisten. Dieses Programm besteht aus drei wesentlichen Bausteinen:
   - einem Datenqualitätsbericht,
   - internen Maßnahmen zur Qualitätskontrolle und
   - einem jährlichen Audit der Daten.

   Nur Daten, die eine vordefinierte Vollständigkeit und Widerspruchsfreiheit erreichen, werden in das Register aufgenommen. Jährlich werden insgesamt 300 bis 625 Datensätze aus 25 zufällig ausgewählten Studienzentren (= 12 bis 25 Datensätze pro Zentrum) auditiert, d.h. die Werte werden mit den Originalwerten verglichen (SDV). Im Jahre 2010 wurden hierbei für die Richtigkeit Werte von 85% bis 97% erreicht.
9. Kommentar
   Das vorgestellte Datenqualitätsprogramm lässt sich für andere Register und Kohorten adaptieren.

### Mphatswe et al. 2012

1. Art des wissenschaftlichen Vorhabens: Intervention zur Verbesserung der Datenqualität
2. Herkunft der Autoren: Südafrika
3. Bezug zum Thema: Feedback, Qualitätsindikatoren, SDV
4. Ort und Art der Intervention: Auf Feedback und Training basierende Intervention zur Verbesserung der Vollständigkeit und Genauigkeit von öffentlichen Routinedaten des Gesundheitswesens in KwaZulu-Natal, Südafrika.
5. Datenbestand: Routinedaten zur primären Gesundheitsversorgung in Südafrika einschließlich Daten zur Prävention einer Mutter-Kind-Übertragung des HIV-Erregers (prevention of mother-to-child transmission [PMTCT]).
6. Qualitätsindikatoren: Genauigkeit, Vollständigkeit.
7. Schlussfolgerungen der Autoren
   Schulung des Pflegepersonals zur Datenaufnahme sowie monatliches Feedback und regelmäßige Prüfung des Datenbestandes konnten die Datenqualität von PMTCT-Daten in Bezug auf Vollständigkeit und Genauigkeit signifikant erhöhen. Demnach spiegelt der Bericht einen einfachen, praktischen Ansatz wider, die Qualität des öffentlichen Gesundheitswesens zu verbessern.
8. Zusammenfassung der Publikation
   Mit 38,7% ist die Prävalenz von HIV im Jahr 2007 bei Frauen in der Region KwaZulu-Natal die höchste in ganz Südafrika. Es wurde postuliert, die Mutter-Kind-Übertragungsrate von HIV bis 2011 auf unter 5% zu senken. Aufgrund der im Vorfeld bekannt gewordenen deutlichen Mängel in der Datenqualität von Routinedaten des Gesundheitswesens, einschließlich Daten zu PMTCT, wurde das Großprojekt 20.000+ Partnership gegründet, welches die Verbesserung der Qualität der PMTCT-Daten in der Provinz KwaZulu-Natal zur Aufgabe hat. Dies beinhaltet die Verbesserung von Vollständigkeit und Genauigkeit der Sammlung von öffentlichen Routinedaten des Gesundheitswesens im District Health Information System (DHIS). Hierzu wurden 78 zufällig gewählte medizinische Einrichtungen der Provinz, die PMTCT-Daten zur Verfügung stellten und an der Intervention teilnahmen, aufgesucht und begleitet. Die Intervention basierte auf drei Hauptmaßnahmen: Training zur Datenerfassung, monatliches Feedback zu den Daten sowie regelmäßige Datenprüfung. Der Datenbestand wurde vor der Intervention zwischen September und November 2007, während der Intervention zwischen Juni und September 2008 sowie nach der Intervention im März 2009 aufgenommen und ausgewertet. Mittels SDV, die den Datenbestand der verschiedenen Einrichtungen mit dem Bestand des DHIS vergleicht, wurden Vollständigkeit und Genauigkeit untersucht und quantifiziert. Dabei verbesserte sich

die Vollständigkeit über alle medizinischen Einrichtungen hinweg von 26% vor der Intervention auf 64% danach. Die Genauigkeit verbesserte sich von 37% auf 65%.

9. Bewertung der Publikation
Die Arbeit bietet ein anschauliches Beispiel für drastische Verbesserungen der Datenqualität durch einfache Maßnahmen wie Schulung und Feedback.

## Nahm et al. 2008

1. Art des wissenschaftlichen Vorhabens: Intervention zur Verbesserung der Datenqualität
2. Herkunft der Autoren: USA
3. Bezug zum Thema: SDV
4. Ort und Art der Intervention: Prüfung einer Verbesserung der Datenqualität bei der Erhebung von Studiendaten mittels Electronic Data Capture (EDC)-Systemen anstelle von papierbasierten Case Report Forms (CRF).
5. Datenbestand: Studiendaten des National Institute of Drug Abuse (NIDA) – Netzwerk für klinische Studien (CTN).
6. Qualitätsindikatoren: nicht anwendbar
7. Schlussfolgerungen der Autoren
In den NIDA-CTN erhobenen Daten ließ sich durch SDV beobachten, dass die Fehlerrate bei der direkten Übertragung von Daten in eine Datenbank, mittels EDC, überraschend gering ist. Analog zur bisherigen Forschung liegt hierbei die Fehlerrate knapp über der bei CRF-/Datenbank-Übertragungen, jedoch stark unter der hier vorkommenden Fehlerrate von Quelldaten-/Datenbank-Übertragungen. Die Autoren weisen jedoch darauf hin, dass aufgrund der Beschaffenheit der CTN-Daten generelle Aussagen zur Thematik nicht möglich sind.
8. Zusammenfassung der Publikation
In den USA hat das NIDA-CTN im Jahr 2005 veranlasst, die Datenqualität von Studien, die EDC Systeme zur direkten Datenübermittlung der Quelldaten an die Studiendatenbank verwenden und dabei eine Prüfung fehlender, unpassender sowie logisch inkonsistenter Werte in den Quelldaten durchführen, zu quantifizieren. Es besteht die Annahme einer besseren Datenqualität gegenüber Datenbanken, deren Quelldaten erst in papierbasierte CRFs und anschließend in die Studiendatenbank übertragen werden. Grund dieser Annahme sind veröffentlichte Fehlerraten, die belegen, dass Fehler vor allem bei der Transkription der Quelldaten in die CRFs, und nicht erst bei der Übertragung der Informationen aus den CRFs in die Datenbanken auftreten. Zur Prüfung der Annahme wurden ab April 2005 vier Studien am Zentrum für Daten und Statistik (DSC) des CTN, die EDC-Systeme verwenden, untersucht. Zu den geprüften Studien wurden Quelldatenabgleiche der Patientendaten zu verschiedenen Zeitpunk-

ten des Studienverlaufs an insgesamt 24 Standorten vorgenommen. Die durchschnittliche Fehlerrate lag hierbei bei 14,3 Fehlern mit einem 95%-Konfidenzintervall von 12–39 Fehlern in jeweils 10.000 Feldern.
Diese Werte sind für Quelldaten-/Datenbank-Übertragungen, verglichen mit veröffentlichten Fehlerraten in Studien die papierbasierte CRFs nutzen, extrem niedrig und vergleichbar mit Fehlerraten von CRF-/Datenbank-Übertragungen. Aufgrund der überraschend niedrigen Ergebnisse wurden die Fehlerraten vier weiterer, ehemaliger, auf CRF basierender CTN-Studien ausgewertet. Die Fehlerraten zwischen Quelldaten und Datenbankinformation dieser Studien waren ebenfalls überraschend niedrig, wobei die Fehlerraten der CRF-/Datenbank-Übertragung meist sogar bei durchschnittlich unter 10 bei 10.000 Feldern lagen. Eine weitere am CTN durchgeführte Studie zu therapeutischen Ansätzen bei Epilepsie mit komplexen medizinischen Daten zeigte jedoch als Ausnahme eine Fehlerrate von durchschnittlich 428 Fehlern in 10.000 Feldern bei direkter Quelldaten-/Datenbank-Übertragung. Dies ist auf den hohen Detailgrad und die hohe notwendige Abstraktion der Daten der Studie zurückzuführen. Erklärt werden die niedrigen Fehlerraten der vorhergehenden CTN-Studien durch die Tatsache, dass es sich bei den datenerfassenden Standorten der CTN-Studien um öffentliche Programme zur Behandlung von Drogenmissbrauch handelt. Die dort erfassten Daten sind oftmals relativ knapp und detailarm. Grundsätzlich gibt es aber auch bei den CTN-Studien folgende Beobachtungen: Die Fehlerraten der CRF-/Datenbank-Übertragung sind sehr niedrig, gefolgt von der durchschnittlichen Fehlerrate der direkten Quelldaten-/Datenbank-Übertragung in Studien die hierzu EDC-Systeme verwenden. Quelldaten-/Datenbank-Fehler die eine CRF als Zwischenstation verwenden sind häufiger. Dies deckt sich mit bisherigen Erkenntnissen.

9. Bewertung der Publikation
   Anhand des Einsatzes von SDV zu Studien, die ETC Systeme zur direkten Datenübertragung der Quelldaten in die Studiendatenbanken verwenden, wurden niedrige Fehlerraten für die Quelldaten-/Datenbank-Übertragung ermittelt. Aufgrund der Einfachheit der Daten, die das CTN erhebt, waren die Fehlerraten der Quelldaten-/Datenbank-Übertragung allerdings auch bei vorausgegangenen Studien, die papierbasierte CRFs verwenden, relativ niedrig. Letztendlich liefert die Arbeit somit nur einen schwachen Beleg für eine Verbesserung der Datenqualität in Studien, die zugunsten der Nutzung von EDC-Systemen, auf papierbasierte CRFs verzichten.

## Sáez et al. 2011

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Spanien
3. Bezug zum Thema: Qualitätsindikatoren

4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: Duplizität, Kontextannotation, Korrektheit, räumliche Stabilität, Vollständigkeit, Vorhersagewert, Widerspruchsfreiheit, Zeitnähe, Zuverlässigkeit
7. Schlussfolgerungen der Autoren
   Basierend auf Charakteristiken, Funktionalitäten und Ergebnissen von sich verändernden Daten wurde ein Datenqualitäts-Vektor (DQV) entworfen. Dieser Vektor soll als allgemeines, anpassbares Werkzeug zur Bewertung von Datenqualität verwendet werden. Das vorgestellte Konzept hat den Anspruch, die Basis weiterführender Arbeiten darzustellen.
8. Zusammenfassung der Publikation
   Mangelnde Datenqualität kann gerade im medizinischen Kontext zu fehlerhaften Entscheidungen mit schwerwiegenden Folgen für den Patienten führen. Falsche Therapieansätze oder Fehlmedikation sind nur einige der Beispiele, die die Autoren in dieser Arbeit aufführen. Bewertung und Handhabung verschiedener Dimensionen der Datenqualität wurden bereits ausführlich in vorhandener Literatur behandelt und diskutiert. Als unzureichend wird allerdings der Umgang mit biomedizinischen Daten, die sich über die Zeit hinweg oder durch die Mehrfachaufnahme an verschiedenen Standorten oder Zentren verändern bzw. unterscheiden können, bewertet. Hierbei sprechen die Autoren vom Problem des „Dataset shift“. Weiterführend wurden deswegen Charakteristiken, Funktionalitäten und Resultate im Bezug auf sich verändernde Daten spezifiziert, die eine Bewertung der Datenqualität erlauben. Als Instrument, mit diesen Eigenschaften und den bereits aus der Literatur bekannten Datenqualitätsdimensionen umzugehen, wird ein DQV präsentiert. Die neun Dimensionen des Vektors wurden wie folgt definiert.
   - Duplizität: Die Feststellung, inwieweit Daten zu einer Entität mehrfach dupliziert erfasst wurden.
   - Kontextannotation: Das Ausmaß der korrekten Annotation von Daten in einem erfassten Kontext.
   - Korrektheit: Das Ausmaß von Genauigkeit und Präzision von Daten in Vergleich zu den tatsächlichen Werten.
   - Räumliche Stabilität: Das Ausmaß der Stabilität von Daten über verschiedene Populationen hinweg.
   - Vollständigkeit: Das Ausmaß der Vollständigkeit der Aufnahme relevanter Daten.
   - Vorhersagewert: Die Beurteilung, inwieweit Daten ausreichend Information für spezifische Entscheidungen beinhalten.
   - Widerspruchsfreiheit: Das Ausmaß, inwieweit Daten spezifizierten Bedingungen und Regeln entsprechen.
   - Zuverlässigkeit: Das Ausmaß an Reputation derjenigen Geschäftsinteressenten und Institutionen, die in die Aufnahme der Daten involviert sind.

- Zeitnähe: Das Ausmaß temporärer Stabilität von Daten.

In der Publikation steht die genaue Spezifikation der Metriken zur Quantifizierung der einzelnen Dimensionen des Vektors noch offen. Einige der Metriken, wie beispielsweise die zur Berechnung der Duplizität, sollen generisch ausfallen, wohingegen andere Dimensionen durch domänenspezifisch angepasste Funktionen berechnet werden sollen. Je nach Anwendung sollen Metriken zu allen oder auch nur Teilmengen der Dimensionen zur Abschätzung der Datenqualität verwendet werden.

9. Bewertung der Publikation
   Wie die Arbeit zeigt, erfordert das Auftreten von sich verändernden Datenelementen neue Herangehensweisen. Zum Umgang mit diesen Elementen stellen die Autoren einen aus neun Dimensionen bestehenden DQV vor. Leider fehlen die konkreten Spezifikationen zur Quantifizierung der einzelnen Dimensionen, weshalb die Ergebnisse wohl vorerst als reine Konzeptvorstellung und nicht als reife Anwendung zu interpretieren sind.

### Salati et al. 2011

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Italien
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Anonymisierte Daten der Datenbank der European Society of Thoracic Surgeons (ESTS) aus dem Zeitraum 2007 bis 2009.
6. Qualitätsindikatoren: Glaubwürdigkeit, Korrektheit/Genauigkeit, Vollständigkeit, Widerspruchsfreiheit.
7. Schlussfolgerungen der Autoren
   Anhand einer Beispieldatenbank aus der Thoraxchirurgie wird in der vorliegenden Studie ein methodischer Ansatz zur Bewertung/Messung verschiedener Datenqualitätskriterien präsentiert. Die Hauptabsicht war dabei nicht die Datenqualitätsanalyse der ESTS-Datenbank, sondern der Test einer Methode, die verallgemeinert als Vorlage auch auf andere medizinische Datenbanken angewendet werden könnte.
8. Zusammenfassung der Publikation
   Bei der Studie geht es um einen methodischen Ansatz, Datenqualität im Bereich der Thoraxchirurgie zu bewerten und eine Vorlage für zukünftige Untersuchungen und Verbesserungen bereitzustellen. Dieser methodische Ansatz wurde zu Testzwecken auf die ESTS-Datenbank angewendet. Bei der ESTS-Datenbank handelt es sich um ein Online-Register aus dem Bereich der Thoraxchirurgie mit ca. 150 Datenfeldern, die die verschiedenen Aspekte unterschiedlicher chirurgischer Prozeduren beschreiben. Hierfür wurden dem Studienteam sämtliche Fälle im Zeitraum Juli 2007 bis Oktober 2009 zur Verfügung gestellt. Die Analyse er-

folgte aus einem ausgewählten Pool von Variablen bei Patienten mit erstgradigem Lungenkrebs, an denen eine Lungenresektion durchgeführt wurde.
Für die Qualitätsmessung wurden folgende fallunabhängige Metriken entwickelt und eingesetzt:
- Glaubwürdigkeit: Der niedrigste Wert von Vollständigkeit, Korrektheit oder Glaubwürdigkeit.
- Korrektheit/Genauigkeit: Anzahl aller korrekten Daten geteilt durch die Anzahl aller gezählten Daten. Bei Abweichung eines Werteintervalls wurden die Daten als inkorrekt klassifiziert.
- Vollständigkeit: Anzahl aller registrierten Daten geteilt durch die Anzahl aller erwarteten Daten.
- Widerspruchsfreiheit: Anzahl aller konsistenten Prüfungen geteilt durch die Anzahl aller Prüfungen. Als inkonsistent wurde beispielsweise der Tod im Krankenhaus bei gleichzeitiger Komplikationsfreiheit bewertet.

Für die ESTS-Datenbank wurde auf den gewählten Datenfeldern eine Vollständigkeit von durchschnittlich 85% berechnet. Die Autoren bewerten dies, mit Ausnahme einiger unterdurchschnittlicher Datenfelder, als akzeptables Ergebnis. Die Korrektheit der Datenbank wurde mit 99% als optimal bewertet. Die Konsistenz beträgt 98%; die Glaubwürdigkeit, die dem niedrigsten Wert der vorhergehenden Qualitätskriterien entspricht, liegt bei 85%. Als Grenzwert für gute Qualität hatte das Studienteam anhand verschiedener Literatursichtungen 80% definiert. Hiermit liegen alle Qualitätskriterien über dem festgelegten Grenzwert.
9. Bewertung der Publikation
   Anschaulich und nachvollziehbar beschreibt die Publikation die Quantifizierung verschiedener Qualitätskriterien anhand einfacher Metriken. Allerdings ist der Grad der Innovation dieser Methodik als eher gering einzuschätzen.

## Shabestari/Roudsari 2013

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Kanada
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: Konkordanz, Korrektheit, Plausibilität, Vollständigkeit, Zeitnähe
7. Schlussfolgerungen der Autoren
   Anhand eines modularen Modells zur Qualitätszusicherung, das auf Erkenntnissen einer Literaturrecherche basiert, kann die Datenqualität von Electronic Health Records (EHRs) verbessert werden.




8. Zusammenfassung der Publikation
   Patientendaten, die in EHRs gespeichert werden, können sowohl in der primären als auch in der sekundären Versorgung von großem Nutzen sein. Wichtig ist hierbei eine hohe Qualität der genutzten Daten, welche sich durch Qualitätskontrolle als auch durch Qualitätssicherung (QS) erreichen lässt. Anhand von Belegen aus der wissenschaftlicher Literatur weisen die Autoren auf gegenwärtige Probleme und Herausforderungen der Datenqualität hin. Im Fokus stehen hierbei die Qualitätskriterien Konkordanz, Korrektheit, Plausibilität, Vollständigkeit und Zeitnähe. Die Vollständigkeit lässt sich generell durch den Vergleich mit einem Goldstandard oder einer Liste erwarteter Elemente prüfen. Zur Gewährleistung der Korrektheit schlagen die Autoren einen Kontrollpunkt direkt bei der Datenerfassung vor. Die Konkordanz lässt sich am besten dann prüfen, wenn Daten aus verschiedenen Datenquellen vorhanden sind. Erreicht werden kann dies beispielsweise durch einen mehrstufigen Aufbewahrungsort der Daten. Für die Prüfung der Plausibilität ist grundsätzlich medizinisches Wissen von Nöten. Die Zeitnähe ist ein komplexes Problem und bedarf einer zentralen Organisation, die die Zeitnähe ihrer Akten kontrolliert. Aus den Erkenntnissen haben die Autoren ein Sequenzdiagramm zu einem fehlervorbeugenden QS-Modell zur Entwicklung von EHRs erstellt. Es besteht aus mehreren verzahnten QS-Modulen und bezieht unter anderem als wichtigen Faktor die Mitarbeit des Pflegepersonals mit ein und basiert demnach nicht ausschließlich auf maschinellen Entscheidungen.
9. Bewertung der Publikation
   In relativ kurzer Form wird von Problemen, Herausforderungen und Lösungsansätzen in Bezug auf verschiedene Kriterien der Datenqualität bei EHRs berichtet. Anhand der Erkenntnisse dieser Literaturrecherche präsentieren die Autoren ein präventives QS-Model. Leider ist die Beschreibung dieses Modells nur relativ oberflächlich; der Nutzen wird nicht weiter validiert.

### Sigurdardottir et al. 2012

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Island
3. Bezug zum Thema: Qualitätsindikatoren, SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: isländisches Krebsregister
6. Qualitätsindikatoren: Aktualität, Validität, Vollzähligkeit
7. Schlussfolgerungen der Autoren
   Das isländische Krebsregister erreicht eine hohe Vollständigkeit, Validität und Aktualität der Daten und ist vergleichbar mit den Krebsregistern der anderen nordischen Staaten.

8. Zusammenfassung der Publikation
   Die Autoren beschreiben die erste formale Evaluation des seit 1954 bestehenden isländischen Krebsregisters. Im Rahmen dieser Evaluation wurden die Vergleichbarkeit, Validität, Aktualität und Vollständigkeit der Daten aus dem Zeitraum 1955 bis 2009 (insgesamt 41.994 Krebsdiagnosen) überprüft. Für die Jahre 2000 und 2001 wurde per Record Linkage zwischen dem Krebsregister und dem Register der Krankenhausentlassungen die Vollzähligkeit des Krebsregisters bestimmt, die bei 99,15% liegt. 96,4% der zwischen 2005 und 2009 registrierten Tumore sind morphologisch verifiziert. Nur 0,2% der Fälle sind DCO (Death Certificate Only)-Fälle. Innerhalb eines Jahres nach Diagnose sind 84,8% der Diagnosen im Krebsregister gemeldet. Insgesamt erreicht das Krebsregister eine hohe Datenqualität.
9. Kommentar
   Dieser Artikel beschreibt eine weitere Anwendung häufig genutzter Datenqualitätsindikatoren.

## Stevens et al. 2008

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Neuseeland
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Durch die stationäre Versorgung erhobene Lungenkrebsdaten der Regionen Auckland und Northland (Neuseeland) aus dem Jahr 2004 sowie Daten des New Zealand Cancer Registry (NZCR) mit übereinstimmenden Einschlusskriterien.
6. Qualitätsindikatoren: Genauigkeit, Vollständigkeit
7. Schlussfolgerungen der Autoren
   Unvollständigkeit als auch Ungenauigkeit der mit Lungenkrebs assoziierten Daten des NZCR könnten sich negativ auf daraus abgeleitete Forschung auswirken. Zwar ist es begrüßenswert, dass das NZCR, als eines der wenigen Register der Welt, nationale Daten zum Krankheitsverlauf führt, jedoch ist das Fehlen dieser Daten in 42% der Fälle trotz meist vorhandener Quelldaten problematisch. Die Erkenntnisse dieser Studie sollten nicht nur im Fall von Lungenkrebs, sondern auch generell für andere Bereiche des NZCR relevant sein.
8. Zusammenfassung der Publikation
   Die Studie befasst sich mit der Prüfung von Vollständigkeit und Genauigkeit der Daten des NZCR. Hierfür wurden rückwirkend alle Patienten aus den Regionen Auckland und Northland, die im Jahre 2004 mit Lungenkrebs diagnostiziert wurden, ermittelt. Dabei konnten 556 Patienten identifiziert und die dazugehörigen Informationen aus den klinischen Akten und regionalen Datenbanken bezogen werden. Bei den Informa-



tionen handelt es sich um die erfassten Daten zu Eingang, Übermittlung und Verwaltung von Lungenkrebspatienten bis zur Initiierung erster Behandlungen oder der Entscheidung, den Patienten ohne Eingriffe weiter zu versorgen. Die Informationen wurden mit denen des NZCR abgeglichen. Lediglich 490 der in den regionalen Datenbanken enthaltenen Patienten waren auch im NZCR gelistet, wobei Informationen zu neun weiteren Patienten mit den passenden Einschlusskriterien ausschließlich im NZCR vorlagen. Die Vollständigkeit des NZCR beträgt hiermit 88% (499/565). Zusätzlich enthielt der bezogene Datensatz des NZCR ein Duplikat sowie 78 (13%) aus verschiedenen Gründen unzulässige Einträge. In den 490 gemeinsamen Einträgen wurden 4 (1%) Unterschiede im Geburtsdatum, 16 (3%) Unterschiede in der Volksabstammung sowie 22 (4%) Unterschiede in den Adressangaben ermittelt. Das Diagnosedatum stimmte nur in 288 (59%) der Fälle überein. Die große Diskrepanz ist jedoch hauptsächlich auf die Unterschiede in den Felddefinitionen der regionalen Datenbanken sowie des NZCR zurückzuführen. Weiterhin unterschieden sich die Einträge in 21 (4%) Fällen bei der Diagnosen-Grundlage sowie 5 (1%) Fällen im Typ des Tumors. Zum Fortschritt des Tumors gab es in 97% der Fälle in den regionalen Datenbanken sowie 58% im NZCR Informationen. Bei 279 gemeinsamen Fällen stimmten dabei die Informationen in 216 (77%) der Fälle überein.

9. Bewertung der Publikation
   Die Studie belegt, dass Datenqualität auch in nationalen Registern Relevanz besitzt. Demnach ist bei der Forschungsarbeit mit Registerdaten grundsätzlich mit einem gewissen Bias und einer statistischen Unsicherheit zu rechnen.

## Taggart et al. 2012

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Australien
3. Bezug zum Thema: Qualitätsindikatoren, Feedback
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Routinemäßiger Bezug von Daten aus Electronic Health Records (EHRs) von vier Arztpraxen der Grundversorgung.
6. Qualitätsindikatoren: Korrektheit, Vollständigkeit, Widerspruchsfreiheit
7. Schlussfolgerungen der Autoren
   Das electronic Practice Based Research Network (ePBRN) kann verwendet werden, um die Datenqualität von Electronic Health Records (EHRs) aus der medizinischen Grundversorgung zu überwachen und zu verbessern. Die Verlinkung der Daten aus verschiedenen Quellen gestattet zudem neue Forschungsanwendungen. Der initiale Fokus des ePBRN liegt auf der Diabetesforschung, soll jedoch ausgeweitet werden.

8. Zusammenfassung der Publikation
Das Paper beschreibt die Vorgehensweise und erste Ergebnisse des ePBRN. Hierbei handelt es sich konzeptionell um ein Data-Warehouse, das regelmäßig Daten von Arztpraxen transportverschlüsselt bezieht, Patientendaten probabilistisch mit Hilfe der Software GRHANITETM abgleicht, zusammenführt und für die Forschung zur Verfügung stellt. Arztpraxen, die mit dem Primary Health Care Research Network (PHReNet) oder der integrierten Pflege des South West Sydney Local Health District Diabetes Service in Kontakt stehen, wurden zur Teilnahme am ePBRN eingeladen. In einer ersten Pilotstudie wurde das System anhand von initial drei, später vier Arztpraxen getestet. Die Auswertung, insbesondere der Qualitätskriterien Vollständigkeit, Korrektheit und Widerspruchsfreiheit, diente dazu, die Datenqualität anhand des Systems zu evaluieren. Die Vollständigkeitsprüfung umfasste zum einen die Prüfung des Vorhandenseins von Datenelementen zu sozialen Faktoren wie Geschlecht, Alter, ethnische Herkunft sowie verschiedener Risikofaktoren wie dem BMI oder Blutdruck. Zum anderen wurde geprüft, ob Daten vorhanden waren, die bei klinischen Entscheidungen im Kontext des Diabetes nutzbar wären. In der Pilotstudie lag die Vollständigkeit von Geschlecht und Geburtsdatum bei allen Praxen bei 100%. Die Vollständigkeit der anderen Daten schwankte je nach Merkmal zwischen 10% und 86%. Dies wurde von den Autoren als zu schwach bewertet, um medizinische Entscheidungen zu treffen. Zur Analyse der Korrektheit wurde geprüft, in welchem Maß Skalen, am Beispiel des Gewichtes die Maßeinheit Kilogramm, bzw. Wertebereiche richtig verwendet wurden. Die Korrektheit lag fast durchgängig bei annähernd 100%. Die Analyse zur Widerspruchsfreiheit verglich die Daten auf einheitliche Terminologien und Kodierungen (z.B. SNOMED). In der Pilotstudie lag die Widerspruchsfreiheit in jedem Attribut bei 100%. Um die Datenqualität zu verbessern, wird den Arztpraxen ein Feedback zur Datenqualitätsauswertung zur Verfügung gestellt.
9. Kommentar
Durch das zentrale Data-Warehouse fallen Datenqualitätsmängel in spezifischen Arztpraxen schnell auf. Das System bietet demnach eine hervorragende Möglichkeit, die Datenqualität zu überwachen und über Rückmeldungen an die Arztpraxen gezielt zu verbessern.

### Thoburn et al. 2007

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: USA
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: An das US-amerikanische National Program of Cancer Registries (NPCR)-Cancer Surveillance System (NPCR-CSS) angeschlossene Krebsregister.

6. Qualitätsindikatoren: Korrektheit, Vollständigkeit
7. Schlussfolgerungen der Autoren
   Für gute Datenqualität in den untersuchten Krebsregistern sind gut qualifizierte Mitarbeiter, der Zugriff auf verschiedene Datenquellen und die Einhaltung von Standards wichtig.
8. Zusammenfassung der Publikation
   Die Daten von 34 Krebsregistern, die an das US-amerikanische National Program of Cancer Registries (NPCR)-Cancer Surveillance System (NPCR-CSS) angeschlossen sind, wurden auf Vollständigkeit und Richtigkeit überprüft. Untersucht wurden 13 Datenelemente aus den Diagnosejahren 1998 bis 2001 (41.512 Tumore). Students t-Test wurde benutzt, um statistisch signifikante Unterschiede bei Covariaten zu finden. Die Vollständigkeit über alle Register hinweg betrug 96,4% und die Korrektheit 95,0%. Eine bessere Datenqualität zeigte sich bei den Registern, die mehr zertifizierte Tumordokumentationskräfte einsetzten und verschiedene Datenquellen nutzten.
9. Kommentar
   Die Untersuchung zeigt die Bedeutung flankierender Maßnahmen für die Erreichung einer guten Datenqualität. Dies lässt sich auch auf andere Register und Kohorten übertragen.

### Tolonen et al. 2006

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Finnland
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Daten zu Individuen aus 38 Populationen im Kontext kardiovaskulärer Erkrankungen.
6. Qualitätsindikatoren: nicht anwendbar
7. Schlussfolgerungen der Autoren
   m MONICA Projekt der WHO werden verschiedene Aspekte der Datenqualität berücksichtigt, dokumentiert und bearbeitet. Die Qualität jedes Datenelements wurde in Qualitätsberichten mithilfe eines Qualitätsscores quantifiziert.
8. Zusammenfassung der Publikation
   In der internationalen Studie WHO MONICA geht es um die Evaluierung von Risikofaktoren zu kardiovaskulären Erkrankungen. Hierfür wurden Daten aus 38 Populationen zu Personen im Alter zwischen 35 und 64 Jahren gesammelt und analysiert. Voraussetzung für eine solche Analyse ist eine zuverlässige Datenqualitätskontrolle, die in dieser Arbeit beschrieben wird.
   Bei der Datenerfassung müssen von den Organisatoren der lokalen Zentren bereits detaillierte Fragebögen zu Quellpopulation, Auswahlmethoden, essmethoden und Messgeräten sowie zum Training des Personals

ausgefüllt werden. Für die Sammlung der Risikofaktoren gibt es ein standardisiertes Protokoll, und die Daten werden durch systematische Fehlerprüfungen als auch detaillierte retrospektive Datenkontrollen geprüft. Hierzu werden in regelmäßigen Abständen Datenqualitätsberichte erstellt und öffentlich zugänglich gemacht. Teil dieser Berichte ist eine Quantifizierung der Datenqualität einzelner Datenelemente anhand eines Datenqualitätsscores. Im Einzelnen handelt es sich bei den untersuchten Datenelementen um das Alter der Teilnehmer, die Quellpopulation, den Überwachungszeitraum sowie die verschiedenen Risikofaktoren. Ein Score von 0 steht für mangelhafte Datenqualität, 1 für geringere Mängel und der Wert 2 für optimale Datenqualität. Quellpopulationen wurden anhand von nationalen oder lokalen Bevölkerungs- und Wahlregistern bestimmt. Der Qualitätsscore hierzu wurde anhand der Anzahl von nicht geeigneten bzw. nicht erreichbaren Individuen berechnet. 20 von 38 der Quellpopulationsbewertungen wurden hierbei mit einem Score von 2 versehen, die restlichen mit einem Wert von 1. Für die erfassten Datenfelder zu Fragestellungen wurde der Score anhand zweier Kriterien berechnet: zum einen der Abweichung zwischen den an den Umfragezentren lokal verwendeten Fragen zu den Fragevorgaben, die zur Verwendung in MONICA empfohlen werden, zum anderen das Vorhandensein von Ergebnisdaten zu jeder Frage. Zur Berechnung des Qualitätsscores der Risikofaktoren wurden zu jeder Messung die verwendeten Instrumente und Prozeduren dokumentiert und mit dem MONICA Protokoll abgeglichen sowie die Ergebnisse der Messungen evaluiert. Weiterführende Untersuchungen zu den Risikofaktoren Cholesterol sowie dem BMI demonstrierten den Zweck der Qualitätsscores in statistischen Arbeiten. Es zeigte sich, dass der BMI in Populationen mit Datenqualitätsproblemen niedriger als bei anderen Populationen ist. Der Quality-Score des BMI kann somit verwendet werden, um Populationen mit niedriger Datenqualität zu erkennen. Zudem kann der Qualitätsscore grundsätzlich als Qualitätsgewicht verwendet werden.

9. Bewertung der Publikation
   Durch die Einführung eines Qualitätsscores bietet die Arbeit eine leicht verständliche Methode, um Datenqualität zu quantifizieren. Dieses Konzept erlaubt erweiterte Möglichkeiten bei der statistischen Auswertung.

## Tuble 2011

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Australien
3. Bezug zum Thema: Qualitätsindikatoren, SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Quelldaten sowie Einträge der Datenbank der Perfusion Downunder Collaboration (PDUC) zu 60 Fällen von kardiopulmonalem Bypass aus den Jahren 2008 bis 2009.




6. Qualitätsindikatoren: Genauigkeit, Vollständigkeit
7. Schlussfolgerungen der Autoren
   Die PDUC-Datenbank ist momentan angemessen akkurat und vollständig. Verbesserungen der Qualitätssicherung sind jedoch denkbar und umfassen das Training von Datenmanagern, die Datenvalidierung sowie ein Feedback zu Datenqualitätsanalysen.
8. Zusammenfassung der Publikation
   Inhalt der Arbeit ist die Analyse der Datenqualität der PDUC-Datenbank. Hierbei handelt es sich um eine Datenbank zum kardiopulmonalen Bypass. Zur Auswertung wurden als Stichprobe Quelldaten von 10% der bearbeiteten Fälle (n = 60) zweier australischer Krankenhäuser aus den Jahren 2008, bzw. die eines weiteren Krankenhauses aus dem Jahr 2009 untersucht und mit den hierzu eingetragenen Daten der PDUC-Datenbank abgeglichen. Insgesamt wurden dabei 57 von 260 möglichen Variablen ausgewertet. Bei den 60 geprüften Fällen wurden insgesamt 3420 Werte geprüft. Hiervon waren 6,9% ungenau übermittelt; 3,2% der Variablenausprägungen fehlten komplett. Weiterhin wurden die Variablen mit Abweichungsraten (0–25%, 25–50%, 51–75%, 75–100%) versehen. Ungenau übermittelte, fehlende, nicht eingegebene oder nicht validierbare Werte wurden als Abweichung gewertet. Es zeigte sich, dass der Großteil der Variablen niedrige Abweichungsraten im Bereich von 0–25% aufwiesen. Höhere Raten waren vor allem auf systematische Unterschiede zurückzuführen. Zufällige Fehler, wie beispielsweise Transkriptionsfehler, traten vor allem bei manuell erfassten Variablen auf.
9. Kommentar
   Mithilfe von SDV gelingt es, die Datenqualität der PDUC-Datenbank zu quantifizieren. Diese ist moderat zufriedenstellend. Zwar ist die Stichprobe mit lediglich 60 Fällen klein, wie die Autoren in der Diskussion der Arbeit beschreiben; die Resultate können jedoch als Anregung für Möglichkeiten zur Verbesserung der Qualitätssicherung von Daten genutzt werden.

### Tudur Smith et al. 2012

1. Art des wissenschaftlichen Vorhabens: Health Technology Assessment
2. Herkunft der Autoren: Vereinigtes Königreich
3. Bezug zum Thema: SDV
4. Ort und Art der Intervention: Im Rahmen einer in Großbritannien durchgeführten Studie über Krebspatienten im fortgeschrittenen Stadium wurden die Effektivität und die Auswirkungen von zentraler als auch standortspezifischer Überwachung mittels SDV geprüft.
5. Datenbestand: Studiendaten zu Nachsorge und Tod von 533 Krebspatienten aus 75 Artzpraxen und Kliniken.
6. Qualitätsindikatoren: nicht anwendbar

7. Schlussfolgerungen der Autoren
   Anhand des Beispiels von Patienten, die unter fortgeschrittenem Krebs leiden, wurde der Nutzen eines lokalen Monitorings mittels SDV mit einem zentralen Monitoring verglichen. Hierbei war der beim lokalen Monitoring zusätzlich betriebene Aufwand hoch und teuer. Die Daten waren nicht zwangsläufig fehlerfrei. Die ermittelten Unterschiede zwischen Studien- und SDV-Daten hatten dabei keinen Einfluss auf die Haupterkenntnisse der Studie.
8. Zusammenfassung der Publikation
   Zwischen Januar 2002 und März 2006 wurden in einer britischen Studie Informationen aus 75 Pflegeeinrichtungen zu 533 Patienten im fortgeschrittenen Krebsstadium erhoben. Die Daten wurden in papierbasierte CRFs aufgenommen, welche die Quelldaten darstellen, und in einer zentralen Studiendatenbank abgespeichert. Durch die unabhängige Erfassung von Geburts- und Todesdaten des Office for National Statistics (ONS) gab es eine parallele zentrale Erfassung von Eintritt, Datum und Ursache des Todes der Studienteilnehmer sowie von deren identifizierenden Daten.
   In den Jahren 2006 und 2007 wurde zusätzlich an den jeweiligen Standorten eine SDV über alle Daten durchgeführt, um zu gewährleisten, dass die in der Studiendatenbank gespeicherten Daten mit denen der CRFs übereinstimmen. Die verifizierten Daten wurden in einer neuangelegten SDV-Datenbank abgelegt. Um den Nutzen der SDV einzuschätzen, wurden die Datenbanken in folgenden Kriterien verglichen:
   - Basisdaten,
   - primäre Folgen (generelles Überleben) sowie
   - sekundäre Folgen (folgenloses Überleben [PFS], Reaktion auf die Maßnahmen [RECIST], schwerwiegende unerwünschten Ereignisse [SAE]).

   Die Unterschiede zwischen den Studien- und den SDV-Daten waren generell niedrig, und die Informationen stellten sich bei einer Sensitivitätsanalyse über alle Daten als nahezu identisch dar. Bei 13 (2,4%) Patienten stimmte das Todesdatum nicht überein. Zu 29 (5,4%) weiteren Patienten lagen Todesdaten vor, die in den Studiendaten nicht gelistet waren. Auch die Kaplan-Meier-Überlebenskurve zeigt kaum Unterschiede zwischen den Studien- und SDV-Daten mit einem vernachlässigbaren Unterschied hinsichtlich der Effektivität der Behandlungen. Zwischen den ONS- und SDV-Daten gab es im Todesdatum 53 (9,9%) Unterschiede, was belegt, dass die SDV-Daten nicht zwangsläufig fehlerfrei sein müssen. Bei 132 (24,8%) der Patienten lieferte der Vergleich der Studien- und SDV-Daten zur PFS-Zeit Unterschiede. Dies hatte jedoch erneut kaum einen Einfluss auf die Kaplan-Meier-Überlebenskurve. Beim Vergleich der RECIST-Klassifikationsdaten waren die Abweichungen größer, sind jedoch hauptsächlich (93,5%) auf Unterschiede in den verwendeten Klas-



sifikationen zurückzuführen. Bei der Anzahl der SAE gab es bei 53 (9,9%) der Patienten Abweichungen.
Der zusätzliche Aufwand des dezentralen Monitorings mit SDV schlug sich für diese Studie in etwa 30,5 Arbeitswochen nieder (bei 7 Stunden am Tag und 5 Werktagen), was von den Autoren mit ca. 25.000 € zusätzlichen Kosten berechnet wurde. Die Kostenschätzung des zentralen Monitorings lag lediglich bei ca. 2.400 €.

9. Bewertung der Publikation
Zwar konnten anhand der SDV zu Krebspatientendaten geringe Abweichungen zwischen Originaldaten und erfassten Daten ermittelt werden, diese hatten jedoch kaum einen Einfluss auf weiterführende Analysen, wie in diesem Beispiel die Auswertung der Kaplan-Meier-Überlebenskurve. Es stellt sich demnach die kritische Frage, ob sich generell die erheblichen Mehrkosten eines dezentralen Monitorings gegenüber einem zentralen Monitoring lohnen.

### Venet et al. 2012

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: Belgien
3. Bezug zum Thema: SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: nicht anwendbar
7. Schlussfolgerungen der Autoren
Zentrale, statistische Überwachung kann abnormale Muster in Daten aufdecken. Hierdurch ist es möglich, lokale Untersuchungen gezielt auf spezifische Studienzentren mit problematischem Datenbestand zu beschränken.
8. Zusammenfassung der Publikation
Im Großteil klinischer Studien wird Datenqualität durch aufwendiges lokales Monitoring und SDV gewährleistet. Um diesen Aufwand und die damit verbundenen Kosten grundlegend zu verringern, wird als Alternative präsentiert, Datenqualität durch ein zentrales, statistisches Monitoring zu verbessern. Beim zentralen, statistischen Monitoring wird der individuelle Datenbestand eines jeden an einer Studie teilnehmenden Zentrums mit dem kollektiven Datenbestand aller Zentren verglichen. Hierdurch ist eine Identifikation abnormaler Datenmuster, wie sie zum Beispiel bei künstlichen oder gefälschten Daten auftreten, möglich. Folglich können nach der Erkennung von Abweichungen gezielte Maßnahmen für einzelne Studienzentren getroffen werden. Das zentrale, statistische Monitoring ist allerdings nur bei ausreichend großen Datenmengen sowie einem Datenbestand aus verschiedenen Datenquellen, beispielsweise verschiedenen Studienzentren, einsetzbar. Weiter-

führend werden auf der Grundlage von echten Studiendaten Beispiele präsentiert, bei denen Qualitätsmängel durch zentrales, statistisches Monitoring erkannt werden konnten, die bei einem lokaler Monitoring trotz gründlichem SDV unerkannt geblieben wären.

9. Bewertung der Publikation
   Die Autoren präsentieren das zentrale, statistische Monitoring als kostensparende Alternative zum Einsatz eines lokalen Monitorings bei allen Studienzentren. Allerdings erscheint die Methode nur für größere Studien und Projekte sinnvoll anwendbar.

### Verhulst et al. 2012

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: Niederlande
3. Bezug zum Thema: SDV
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Daten des internationalen Mucopolysaccharidose-Typ I-Registers
6. Qualitätsindikatoren: Korrektheit
7. Schlussfolgerungen der Autoren
   Die globalen Fehlerraten zeigen eine akzeptable Datenqualität in den untersuchten Variablen des Registers.
8. Zusammenfassung der Publikation
   Das internationale Mucopolysaccharidose (MPS)-Typ I-Register wurde aufgrund regulatorischer Anforderungen im Jahre 2003 angelegt und sammelt auf freiwilliger Basis Informationen zum Krankheitsverlauf bei Patienten mit MPS Typ I. Die gesammelten Daten werden mittels Queries, monatlichen Reviews und elektronischen Audits auf fehlende, inkonsistente und nicht-valide Daten überprüft. Mittels Originaldatenabgleich (SDV) wurde die Korrektheit der Registerdaten festgestellt. Die SDV wurde in ausgewählten Zentren in Europa, Süd-Amerika und Nord-Amerika durchgeführt. Jeweils 3 Patienten pro Zentrum wurden zufällig ausgewählt und die wichtigsten Variablen zur Beschreibung der Erkrankung und ihrer Behandlung zu Baseline und der letzten verfügbaren Untersuchung betrachtet. Die SDV wurde in den Jahren 2008 und 2009 (Phase 1) sowie 2009 und 2010 (Phase 2) durchgeführt. In Phase 1 zeigte sich eine globale Fehlerrate von 2,7%; in Phase 2 stieg die Rate auf 3,7%. Systematische Fehler wurden nicht gefunden.
9. Kommentar
   Dieser Artikel zeigt die Sinnhaftigkeit des Originaldatenabgleichs (SDV).

### Weiskopf/Weng 2013

1. Art des wissenschaftlichen Vorhabens: systematisches Review
2. Herkunft der Autoren: USA

3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: nicht anwendbar
6. Qualitätsindikatoren: Konkordanz, Korrektheit, Plausibilität, Vollständigkeit, Zeitnähe
7. Schlussfolgerungen der Autoren
   Die in der Literatur beschriebenen Qualitätsindikatoren ließen sich grob in fünf Kategorien zusammenfassen. Hiervon bewerten die Autoren nur drei dieser Kategorien als fundamental (Korrektheit, Vollständigkeit, Zeitnähe). Zudem bemängeln die Autoren generell die inkonsistente Terminologie bei der Diskussion um Datenqualität und vermerken die teilweise große Varianz in den Ergebnissen der Qualitätsabschätzungen. Sollte das Bestreben vorhanden sein, Daten aus Electronic Health Records (EHRs) als zukünftige Quelle medizinischer Forschung zu nutzen, müssen validierte, systematische Methoden für die EHR-Qualitätsbewertung entwickelt werden.
8. Zusammenfassung der Publikation
   Bei der Arbeit handelt es sich um ein systematisches Review zur Datenqualitätsbewertung in Bezug auf die Wiederverwendung von EHR-Daten für die Forschung. Auf PubMed wurden mit Schlagworten, die Indikatoren von Datenqualität beschreiben, sowie aus verschiedenen, zu EHR synonymen MeSH-Terms, insgesamt 230 Literaturquellen ermittelt (Juli 2012). Die Suchanfrage lautete wie folgt: ((‘data quality’ OR ‘data accuracy’ OR ‘data reliability’ OR ‘data validity’ OR ‘data consistency’ OR ‘data completeness’ OR ‘data errors’ OR ‘data error’) AND (EHR OR electronic medical record OR computerized medical record OR medical records systems, computerized [mh]) AND English[lang])). Nach manueller Prüfung auf verschiedene Einschlusskriterien wurden 95 Artikel für die weitere Bearbeitung ausgewählt. Zu diesen Artikeln wurden die darin untersuchten Qualitätsindikatoren, die Bewertungsmethoden und der untersuchte Datenbestand ausgewertet. Insgesamt konnten 27 spezifische Qualitätsindikatoren ermittelt werden. Diese ließen sich wie folgt gruppieren:
   - Vollständigkeit (in 64% der Artikel),
   - Korrektheit (in 60% der Artikel),
   - Konkordanz (in 17% der Artikel),
   - Plausibilität (in 7% der Artikel) und
   - Zeitnähe (in 4% der Artikel)

   Zusätzlich konnten sämtliche Methoden zur Datenqualitätsbewertung in die folgenden sieben Kategorien unterteilt werden: Goldstandard, Übereinstimmen von Datenelementen, Vorhandensein von Elementen, Übereinstimmen von Quelldaten, Vergleich der Verteilungen, Validitätscheck und Log Review. Der Großteil der Studien befasste sich zudem mit strukturierten Daten (73%).

9. Bewertung der Publikation
   Das Review gibt einen ausführlichen Überblick über genutzte Qualitätsindikatoren und deren Häufigkeit im Kontext der Wiederverwendung von EHR-Daten für die Forschung. Dabei zeigte sich, dass vor allem Vollständigkeits- und Korrektheitsanalysen zentrale Punkte bei der Bewertung von Datenqualität sind.

### Wu et al. 2008

1. Art des wissenschaftlichen Vorhabens: Konzeptvorstellung
2. Herkunft der Autoren: USA
3. Bezug zum Thema: Qualitätsindikatoren
4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Patienten mit Koronararterien-Bypass
6. Qualitätsindikatoren: Rekrutierungsrate, Vollzähligkeit
7. Schlussfolgerungen der Autoren
   Das von den Autoren vorgeschlagene modifizierte Maß zur Schätzung der Vollzähligkeit eines Follow-Ups liegt bei der betrachteten Population näher an der Realität als das Original-Maß.
8. Zusammenfassung der Publikation
   Die Autoren Clark, Altman und De Stavola [Clark et al. 2002] haben ein Maß „C“ für die Vollzähligkeit einer Nachverfolgung von Patienten (Follow-Up) vorgeschlagen, das als Prozentsatz der aktuell beobachteten Follow-Up-Jahre bezogen auf das Maximum möglicher Follow-Up-Jahre definiert ist. Die Autoren des vorliegenden Artikels bemängeln, dass „C“ die Vollzähligkeit eines Follow-Ups unterschätzt, da der Nenner keine unbeobachteten Todesfälle berücksichtigt. Daher schlagen die Autoren eine Modifikation „C*“ vor, die dieses Problem berücksichtigt. Dieser Ansatz wurde mit den Daten eines Follow-Ups einer Gruppe von Patienten mit Koronararterien-Bypass getestet. Zur Ermittlung der tatsächlichen Vollzähligkeit wurde der Nationale Todesindex der USA herangezogen. Die tatsächliche Vollzähligkeit lag bei 85,0%. Das modifizierte „C*“ lag mit 84,5% näher an der Realität als das Original-„C“ mit 80,4%.
9. Kommentar
   Dies ist ein interessanter Ansatz zur Schätzung der Vollzähligkeit eines Follow-Ups, und er scheint auch für Register und Kohorten geeignet zu sein. Eine umfassende Evaluierung dieses Ansatzes wäre wünschenswert.

### Xian et al. 2012

1. Art des wissenschaftlichen Vorhabens: Analyse von Datenqualität
2. Herkunft der Autoren: USA
3. Bezug zum Thema: Qualitätsindikatoren




4. Ort und Art der Intervention: nicht anwendbar
5. Datenbestand: Quelldokumente zu 438 Schlaganfallpatienten und den dazugehörigen Einträgen in das Get With The Guidelines (GWTG) Stroke-Register.
6. Qualitätsindikatoren: Genauigkeit, Zuverlässigkeit.
7. Schlussfolgerungen der Autoren
   Die in dieser Studie durchgeführte Validierung von zufällig ausgewählten Daten des GWTG Stroke-Registers sowie den dazugehörigen Quelldokumenten der jeweiligen Krankenhäuser bestätigt die generell hohe Datenqualität der GWTG Stroke-Datenbank.
8. Zusammenfassung der Publikation
   Zwischen Oktober 2008 und September 2009 wurden 223 (von insgesamt 1342) Krankenhäuser, die regelmäßig an das GWTG Stroke-Register Daten übermitteln, zufällig ausgewählt, um an einer Qualitätsanalyse des Datenbestandes des GWTG Stroke-Registers teilzunehmen. Letztendlich beteiligten sich an dieser freiwilligen Studie 147 Krankenhäuser mit 438 zur Verfügung gestellten medizinischen Quelldokumenten. Diese wurden hierbei an Outcome Science, Inc. verschickt. Die Firma bietet ein Patienten-Management-Tool an, mit dem die Daten für das GWTG Stroke-Register routinemäßig erfasst werden. Von identifizierenden Merkmalen befreit wurden die abstrahierten Quelldokumente mit den dazugehörigen, in das GWTG Stroke-Register eingetragenen Daten zur weiteren Auswertung an das Duke Clinical Research Institute weitergeleitet. Die Datenauswertung erfolgte durch geschulte Prüfer, die ein standardisiertes Datenauswertungsinstrument, das speziell für die Studie entworfen wurde, verwendeten. Als akkurat wurden nur die Datenelemente bewertet, die sowohl in den Daten der abstrahierten Quelldokumente der Krankenhäuser als auch denen der erfassten GWTG-Daten exakt übereinstimmten. Hierbei konnten 4% der durch das GWTG Stroke-Register aufgenommenen Daten nicht durch Informationen in den abstrahierten Quelldokumenten identifiziert werden. Über alle Daten hinweg betrug die Genauigkeit je nach Datenfeld zwischen 79,0% und 99,5%, durchschnittlich 96,1%. Zudem wurden zeitsensitive Datenfelder auf ihre Zuverlässigkeit geprüft. Hierbei gab es in insgesamt 38 Fällen Unterschiede. Insgesamt sind die Ergebnisse als gut zu bewerten mit Ausnahme einiger Datenfelder, insbesondere Gewichts- und Datumsangaben. Ein Nebenbefund war, dass Daten von akademischen Lehrkrankenhäusern eine etwas höhere Genauigkeit aufwiesen als die der anderen teilnehmenden Krankenhäuser.
9. Bewertung der Publikation
   Die Datenqualität des in der Studie vorgestellten nationalen Registers erscheint zufriedenstellend. Ein Grund hierfür ist wohl die Art der Datenerfassung, die in allen teilnehmenden Zentren über ein standardisiertes Internettool einer externen Firma erfolgte.

# Weiterführende Online-Ressourcen

Unter www.tmf-ev.de/Leitlinie-Datenqualitaet stellt die TMF weiterführende Materialien und Ressourcen zur Leitlinie Datenqualität bereit und macht Inhalte der ersten Auflage verfügbar.

## Literatursichtung bis 2005/2006

Grundlage von Version 1.0 der Leitlinie bildeten eine umfassende Literatursichtung und eine Befragung von ausgewiesenen Experten, die der 1. Auflage des Buches als Anhang beigefügt war. Die Zusammenfassung von Literatursichtung und Expertenbefragung bis 2005/2006 steht nun unter www.tmf-ev.de/Leitlinie-Datenqualitaet digital zur Verfügung.

## Schlussbericht zu TMF-Projekt V020-04

Im Vorlauf zur Überarbeitung der Leitlinie wurden neue Indikatoren im Rahmen des TMF-Projektes V020-04 „Leitlinie Datenqualität als Rahmen für empirische Forschungsvorhaben“ identifiziert. Der Schlussbericht zu diesem Projekt ist unter www.tmf-ev.de/Leitlinie-Datenqualitaet verlinkt.

## Checklisten

Die Checklisten, die als Implementierungshilfen zu Version 1.0 der Leitlinie Datenqualität entwickelt wurden, können unter www.tmf-ev.de/Leitlinie-Datenqualitaet im PDF-Format abgerufen werden.

## Statistische Tabellen und Software

Version 1.0 der Leitlinie wurde darüber hinaus durch statistische Tabellen (www.tmf-ev.de/Produkte/P020021) und spezifische Software (www.tmf-ev.de/Produkte/P020031) ergänzt, die weiterhin im Produktbereich der TMF-Website zur Verfügung stehen und unter www.tmf-ev.de/Leitlinie-Datenqualitaet verlinkt sind.

TMF – Forscher vernetzen
Lösungen bereitstellen
Doppelarbeit vermeiden

## Die TMF sorgt für Qualitäts- und Effizienzsteigerung in der medizinischen Forschung

Die moderne medizinische Forschung steht vor zunehmend komplexen Herausforderungen, für deren Lösung sich die Akteure aus Grundlagenforschung, klinischer Forschung, Versorgungseinrichtungen, Industrie und weiteren Partnern miteinander vernetzen und gemeinsame Strategien entwickeln müssen. Ein zentraler Ansatz ist die Effizienzsteigerung auf allen Ebenen der medizinischen Forschungs- und Entwicklungskette, um – bei gesicherter Qualität – Forschungsergebnisse auf schnellstem Wege in die Patientenversorgung zu übertragen und damit zu einem effizienten und leistungsfähigen Gesundheitswesen beizutragen. Die Bundesregierung unterstützt diesen Prozess unter anderem im Rahmen des Gesundheitsforschungsprogramms und fördert seit mehr als zehn Jahren konsequent die medizinische Verbundforschung. Erfolgreiche Beispiele sind die herausragenden Ergebnisse aus den Kompetenznetzen in der Medizin oder den Koordinierungszentren für Klinische Studien.

Die TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung (kurz: TMF), die vom Bundesministerium für Bildung und Forschung (BMBF) gefördert wird, leistet hierzu einen entscheidenden Beitrag, indem sie Forscher disziplinübergreifend zusammenbringt und Lösungen für die vernetzte medizinische Forschung bereitstellt. Damit übernimmt sie eine wesentliche nationale Aufgabe zur Qualitäts- und Effizienzsteigerung für die Forschung.

### Ziele und Aufgaben

Als Dachorganisation für die medizinische Verbundforschung verfolgt die TMF das Ziel, die organisatorischen, rechtlichen-ethischen und technologischen Voraussetzungen für die klinische, epidemiologische und translationale Forschung zu verbessern. Sie hat die Aufgabe, die wissenschaftliche Arbeit der modernen medizinischen Forschung, die heutzutage überwiegend in kooperativen Projekten mit mehreren beteiligten Standorten stattfindet, zu unterstützen. Dazu stellt sie – öffentlich und gemeinfrei, also für jeden Forscher nutzbar – Gutachten, generische Konzepte, Leitfäden und IT-Anwendungen ebenso wie Schulungs- und Beratungsangebote bereit. Der überwiegende Teil der Produkte steht unter www.tmf-ev.de zum Download zur Verfügung. Ausgewählte Ergebnisse werden in der Schriftenreihe der TMF publiziert.

Die Produkte werden – von der Forschung für die Forschung – von den Fachexperten der Mitgliedsverbünde entwickelt, die in den interdisziplinären Arbeitsgruppen der TMF zusammenkommen. Als Grundmuster und Leitmotiv der gemeinsamen Arbeit in den Arbeitsgruppe gilt der Anspruch, gemeinsame Probleme gemeinsam zu lösen, von vorhandenen Erfahrungen gegenseitig zu profitieren, Doppelarbeit zu vermeiden sowie professionelle Lösungen zu erarbeiten, zu diesen einen Konsens in der Forschergemeinschaft herzustellen und ihre konsequente Nutzung und langfristige Verfügbarkeit zu gewährleisten.

## Geschichte

Die TMF wurde 1999 unter dem Namen „Telematikplattform für Medizinische Forschungsnetze“ als Förderprojekt des BMBF gegründet. Mit dem Ziel, die Struktur zu verstetigen und die gemeinsame Querschnittseinrichtung der medizinischen Verbundforschung noch stärker in die Hände der Forscher selbst zu legen, wurde 2003 der TMF e.V. gegründet. Seither ist die Zahl der Mitgliedsverbünde stark angewachsen. Damit zusammenhängend hat sich auch das thematische Spektrum der TMF verbreitert, die zunächst primär auf Fragen der IT-Infrastruktur ausgerichtet war. Die Themen reichen heute von rechtlichen und ethischen Rahmenbedingungen und Fragen der IT-Infrastruktur über Qualitätsmanagement und Standards für klinische Studien sowie den Themenkomplex Biobanken und molekulare Medizin bis hin zum Problem der Verzahnung von Forschung und Versorgung oder Fragen der Verbundkoordination und der Wissenschaftskommunikation.

2010 beschloss die Mitgliederversammlung eine Umbenennung der TMF, da der Begriff „Telematikplattform“ diesem breiten Spektrum nicht mehr gerecht wurde. Der seither geführte Name „TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V.“ erfasst die Aufgaben und Themen der TMF auf spezifischere Weise.

## Mitglieder

Mitglieder der TMF sind überregionale medizinische Forschungsverbünde, vernetzt arbeitende universitäre und außeruniversitäre Forschungsinstitute, Methodenzentren, regionale Verbundprojekte sowie kooperative Studiengruppen. Dazu gehören unter anderem

- die Deutschen Zentren der Gesundheitsforschung,
- die Kompetenznetze in der Medizin,
- die Koordinierungszentren bzw. Zentren für Klinische Studien (KKS/ZKS),
- die Integrierten Forschungs- und Behandlungszentren,
- die Netzwerke für Seltene Erkrankungen,
- die Fraunhofer-Gesellschaft (mit dem Fraunhofer ITEM als direktem Mitglied),
- die Zoonosen-Forschungsverbünde,
- die zentralisierten Biomaterialbanken
- Universitätsinstitute,
- Patientenorganisationen
- und zahlreiche weitere.

Über Mitgliedsverbünde sind bundesweit alle Universitätsklinika und zahlreiche außeruniversitäre Forschungsstandorte in unterschiedlicher Weise in die TMF eingebunden. Mit Kooperationspartnerschaften sorgt die TMF auch darüber hinaus für eine Einbindung der relevanten Institutionen im Gesundheitswesen.

## Themen und Arbeitsweise

Die durch die Forschungsverbünde und -einrichtungen gemeinsam zu bearbeitenden Querschnittsaufgaben gehen weit über Fragen von Informations- und Kommunikationstechnologie im technischen Sinne hinaus. Die Wissenschaftler in den Forschungsprojekten brauchen Unterstützung und Erfahrungsaustausch in großer Breite:

- zu Fragen der konkreten Umsetzung von Datenschutz und ethischen Richtlinien,
- zum Aufbau von Forschungsinfrastrukturen wie Datenbanken für Forschungsregister und Biobanken,
- zur strategischen Nutzung von Informationstechnologie für die Prozessunterstützung wie für die wissenschaftliche Auswertung,
- zu Rechtsfragen in vielerlei Hinsicht, beispielsweise zum Vertragsrecht innerhalb von Netzwerken, zu Patienteneinwilligungen oder zu Verwertungsfragen,
- zu Fragen der Organisation und des Managements von Forschungsnetzen und ihren Projekten sowie
- zunehmend auch zu Fragen des Budgetmanagements, der Finanzierung und der Nachhaltigkeit von mit öffentlichen Geldern aufgebauten Netzwerkstrukturen.

Alle diese Fragen werden kontinuierlich in den Arbeitsgruppen der TMF bearbeitet, in denen sich die jeweiligen Fachleute aus den verschiedenen Projekten und Forschungsstandorten interdisziplinär zusammenfinden. Dabei entstehen strategische Anstöße und Impulse für die Forschungsinfrastruktur, vor allem aber konkrete Hilfen, Produkte und Services für den Forscher. Regelmäßig tagen einzelne Arbeitsgruppen auch gemeinsam, um auf diese Weise auch themenübergreifende Aspekte aufnehmen und Doppelaktivitäten der Arbeitsgruppen vermeiden zu können.

## Arbeitsgruppen

Die Arbeitsgruppen initiieren Projekte und betreuen sie im Verlauf – bis hin zur Implementierung der Ergebnisse und zur Beratung von Forschungsprojekten auf dieser Basis. Neue Projektvorschläge durchlaufen ein mehrstufiges Auswahlverfahren – von der fachlichen Prüfung und Schärfung in den Arbeitsgruppen über Beratung in der Geschäftsstelle bis hin zur Begutachtung durch den Vorstand. Mit diesem Vorgehen wird sichergestellt, dass die in den Projekten adressierten Probleme für die Forschergemeinschaft relevant sind und dass die angestrebte Lösung einen breiten Konsens für die spätere Anwendung findet.

Arbeitsgruppen können in der TMF je nach aktuellem Bedarf neu eingerichtet, zusammengelegt oder auch aufgelöst werden, wenn ein Thema keine hohe Relevanz mehr hat. Derzeit sind neun Arbeitsgruppen aktiv:

- Arbeitsgruppe Datenschutz
- Arbeitsgruppe IT-Infrastruktur und Qualitätsmanagement
- Arbeitsgruppe Biomaterialbanken
- Arbeitsgruppe Molekulare Medizin
- Arbeitsgruppe Management Klinischer Studien
- Arbeitsgruppe Medizintechnik
- Arbeitsgruppe Zoonosen und Infektionsforschung
- Arbeitsgruppe Netzwerkkoordination
- Arbeitsgruppe Wissenschaftskommunikation

Der interdisziplinäre Austausch wird über die Arbeitsgruppen hinaus durch zahlreiche Symposien und Workshops, durch den TMF-Jahreskongress sowie durch Foren – aktuell beispielsweise zum Thema Versorgungsforschung – ergänzt.

### Lösungen stehen frei zur Verfügung

Die TMF stellt Gutachten, generische Konzepte, Leitfäden und IT-Anwendungen ebenso bereit wie sie Schulungs- und Beratungsservices der Arbeitsgruppen, auch in Form von Einzelberatungen, anbietet. Die Ergebnisse der Arbeit in der TMF stehen öffentlich und gemeinfrei zur Verfügung.

Mit diesem offenen Ansatz verfolgt die TMF das Ziel,

- methodisches Know-how und Infrastrukturen für die vernetzte medizinische Forschung breit verfügbar zu machen,
- die Harmonisierung, die Interoperabilität und das Qualitätsmanagement in der vernetzten medizinischen Forschung durch entsprechende Infrastruktur, Leitfäden und Services zu stärken,
- die Kollaboration in der deutschen medizinischen Forschung sowie deutsche Forscher in internationalen Kooperationen zu stärken,
- die Verstetigung und Nachhaltigkeit akademischer medizinischer Forschungsprojekte zu unterstützen und
- einen Beitrag zu sinnvollem Mitteleinsatz in der öffentlich geförderten medizinischen Forschung zu leisten, indem sie Doppelentwicklungen vermeiden hilft und die Wiederverwendung vorhandener Lösungen organisiert.

Mit ihren Lösungen adressiert die TMF vor allem die nicht-kommerzielle, akademische – universitäre wie außeruniversitäre – Forschung in Deutschland. Unabhängig davon ist aber auch ein steigendes Interesse an den Angeboten aus der Industrie zu verzeichnen. Viele Lösungen der TMF sind zudem auch für das Ausland, insbesondere die deutschsprachigen Länder, relevant und werden in dortigen Forschungseinrichtungen bereits genutzt.

Alle Download-geeigneten Produkte und Ergebnisse stehen auf der TMF-Website zur Verfügung. Einzelne Software-Werkzeuge sind sehr komplex und

bedürfen einer individuellen Anpassung und Erläuterung, so dass sie nur über den direkten Kontakt zur TMF-Geschäftsstelle erhältlich sind, die dann auch für die Betreuung bei der Implementierung und Nutzung des Produktes sorgt. Darüber hinaus fließen die Ergebnisse kontinuierlich auch in die Diskussionen in den Arbeits- und Projektgruppen ein, und sie werden in konkreten Beratungsgesprächen sowie in Schulungs- und Informationsveranstaltungen vermittelt.

### TMF-Schriftenreihe

Wichtige Konzepte, Leitfäden und Hilfstexte veröffentlicht die TMF in ihrer Schriftenreihe, die sie seit mehreren Jahren bei der Medizinisch Wissenschaftlichen Verlagsgesellschaft herausgibt. So erschienen 2006 als erster Band die generischen Lösungen zum Datenschutz für die Forschungsnetze in Buchform (Reng et al.: Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin, Berlin 2006 – Bd. 1). In der Zwischenzeit sind diese Konzepte einer grundlegenden Revision unterzogen und erneut mit den Bundes- und Landesdatenschützern abgestimmt worden. Die überarbeiteten Konzepte werden in Kürze ebenfalls in der TMF-Schriftenreihe publiziert.

Es folgte das Rechtsgutachten zum Aufbau und Betrieb von Biomaterialbanken (Simon et al.: Biomaterialbanken – Rechtliche Rahmenbedingungen, Berlin 2006 – Bd. 2), das im Februar 2008 um einen weiteren Band zum Thema Qualitätssicherung von Biobanken ergänzt wurde (Kiehntopf/Böer: Biomaterialbanken – Checkliste zur Qualitätssicherung, Berlin 2008 – Bd. 5). Das Datenschutzkonzept, das ursprünglich als Bd. 6 der Schriftenreihe publiziert werden sollte, wird in die Veröffentlichung der neuen generischen Datenschutzkonzepte der TMF integriert.

Mit der Checkliste zur Patienteneinwilligung legte die TMF Ende 2006 ein Referenzwerk vor, das den Anwendern ermöglicht, auf der Basis von relevanten, dokumentierten und kommentierten Quellen Patienteninformationen und Einwilligungserklärungen für klinische Studien zu erstellen, die den regulatorischen Anforderungen entsprechen (Harnischmacher et al.: Checkliste und Leitfaden zur Patienteneinwilligung, Berlin 2006 – Bd. 3). Wie die meisten anderen Buchpublikationen auch, wird dieser Band durch weitere online verfügbare Materialien (z.B. Musterverträge) oder Services ergänzt.

An die TMF-Ergebnisse im Bereich Datenschutz und Patienteneinwilligung knüpft auch der 2012 erschienene Band 10 an (Goebel/Scheller: Einwilligungserklärung und Forschungsinformation zur Gewinnung tierischer Proben, Berlin 2012 – Bd. 10). Die Ergebnisse sind im Auftrag der Nationalen Forschungsplattform für Zoonosen erarbeitet worden. Sie dienen dazu, Forschenden Rechtssicherheit bei der Entnahme und Bearbeitung von Tierproben zu geben und sie bei der Erstellung der relevanten Einwilligungsunterlagen zu unterstützen.

Ein Rechtsgutachten zum Problemfeld der Verwertungsrechte in der medizinischen Forschung (Goebel/Scheller: Verwertungsrechte in der medizinischen Forschung, Berlin 2008 – Bd. 7) erschien 2008 als erste Veröffentlichung einer Reihe von Rechtsgutachten, die die TMF zu verschiedenen Fragen hat erstellen lassen, unter anderem zum Thema „elektronische Archivierung von Studienunterlagen". Die Publikation dieser weiteren Rechtsgutachten in der TMF-Schriftenreihe wird sukzessive folgen.

Mit Band 8 (Mildner [Hrsg]: Regulatorische Anforderungen an Medizinprodukte, Berlin 2011 – Bd. 8) legte die TMF erneut die Aufarbeitung eines im Umbruch befindlichen Feldes vor. Das Buch bietet eine Einführung in den regulatorischen Prozess bei der Entwicklung von Medizinprodukten und stellt Handlungshilfen bereit. Dabei wird der gesamte Bereich von der klinischen Bewertung bis zum Health Technology Assessment abgedeckt.

Praktische Empfehlungen für die Verarbeitung und Analyse von Daten, die bei der Hochdurchsatz-Genotypisierung anfallen, gibt Band 9 (Krawczak/Freudigmann [Hrsg.]: Qualitätsmanagement von Hochdurchsatz-Genotypisierungsdaten, Berlin 2011 – Bd. 9), der ebenfalls 2011 publiziert werden konnte. Dabei reichen die behandelten Fragen von Problemen der Validität und Plausibilität über die Erkennung und Vermeidung von Fehlern bis hin zu Anforderungen an Datenhaltung und Datentransfer.

Das vorliegende jüngste Buch ist die Neuauflage des vierten Bandes der TMF-Schriftenreihe, der bereits 2007 erstmals erschienen ist und jetzt in einer aktualisierten und ergänzten Fassung vorgelegt wird. Die Leitlinie (Nonnemacher et al.: Datenqualität in der medizinischen Forschung, Berlin 2014 – Bd. 4) enthält Empfehlungen zum Management von Datenqualität in Registern, Kohortenstudien und Data Repositories.

### Weitere Informationen und Kontakt

TMF – Technologie- und Methodenplattform
für die vernetzte medizinische Forschung e.V.
Charlottenstraße 42/Ecke Dorotheenstraße
10117 Berlin
Tel.: 030 – 22 00 24 7-0
Fax: 030 – 22 00 24 7-99
E-Mail: info@tmf-ev.de
Internet: www.tmf-ev.de

## Zur Schriftenreihe der TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e. V.

In der TMF – Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V. haben sich Netzwerke und vernetzt arbeitende Einrichtungen zusammengeschlossen, um gemeinsam die Fragestellungen und Herausforderungen von medizinischer Forschung an verteilten Standorten zu lösen und die Erfahrungen zu bündeln. Durch den Community-Ansatz erfahren die Ergebnisse der TMF eine breite inhaltliche Abstimmung in der medizinischen und medizininformatisch-biometrischen Fachwelt. Mit ihrer Schriftenreihe macht die TMF die Lösungen einer breiteren Leserschaft zugänglich.

**Bisher in der Schriftenreihe erschienen:**

*Band 1:*
**Generische Lösungen zum Datenschutz für die Forschungsnetze in der Medizin**
von Carl-Michael Reng | Peter Debold
Christof Specker | Klaus Pommerening
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2006

*Band 2:*
**Biomaterialbanken – Rechtliche Rahmenbedingungen**
von Jürgen Simon | Rainer Paslack | Jürgen Robienski
Jürgen W. Goebel | Michael Krawczak
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2006

*Band 3:*
**Checkliste und Leitfaden zur Patienteneinwilligung Grundlagen und Anleitung für die klinische Forschung**
von Urs Harnischmacher | Peter Ihle | Bettina Berger
Jürgen Goebel | Jürgen Scheller
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2006

*Band 4:*
**Datenqualität in der medizinischen Forschung**
von Michael Nonnemacher | Dorothea Weiland
Jürgen Stausberg
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2007

*Band 5:*
**Biomaterialbanken – Checkliste zur Qualitätssicherung**
von Michael Kiehntopf | Klas Böer
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2008

*Band 7:*
**Verwertungsrechte in der vernetzten medizinischen Forschung**
von Jürgen W. Goebel | Jürgen Scheller
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2009

*Band 8:*
**Regulatorische Anforderungen an Medizinprodukte**
von Kurt Becker | Sandra Börger | Horst Frankenberger
Dagmar Lühmann | Thomas Norgall
Christian Ohmann | Annika Ranke | Reinhard Vonthein
Andreas Ziegler | Andreas Zimolong
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2011

*Band 9:*
**Qualitätsmanagement von Hochdurchsatz-Genotypisierungsdaten**
von Michael Krawczak | Mathias Freudigmann (Hrsg.)
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2011

*Band 10:*
**Einwilligungserklärung und Forschungsinformation zur Gewinnung tierischer Proben**
von Jürgen W. Goebel | Jürgen Scheller
MWV Medizinisch Wissenschaftliche Verlagsgesellschaft, 2012